

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

87 931154

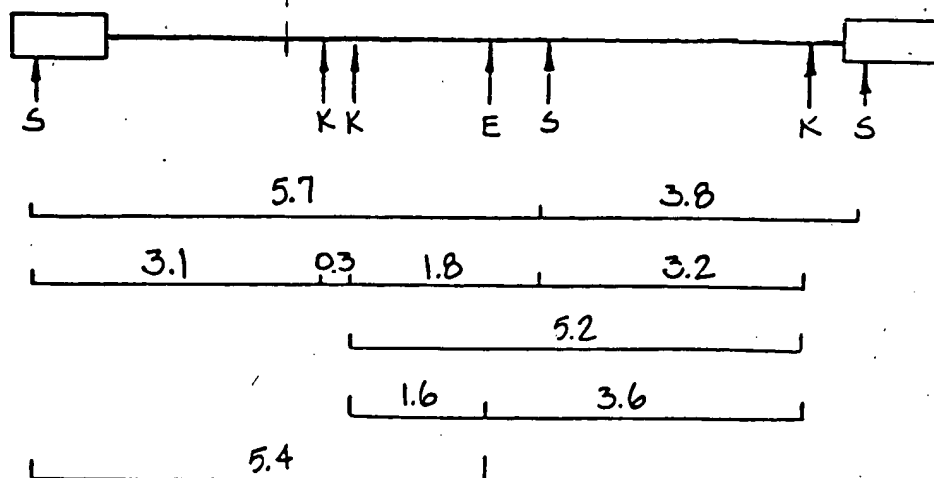


FIG. 1.

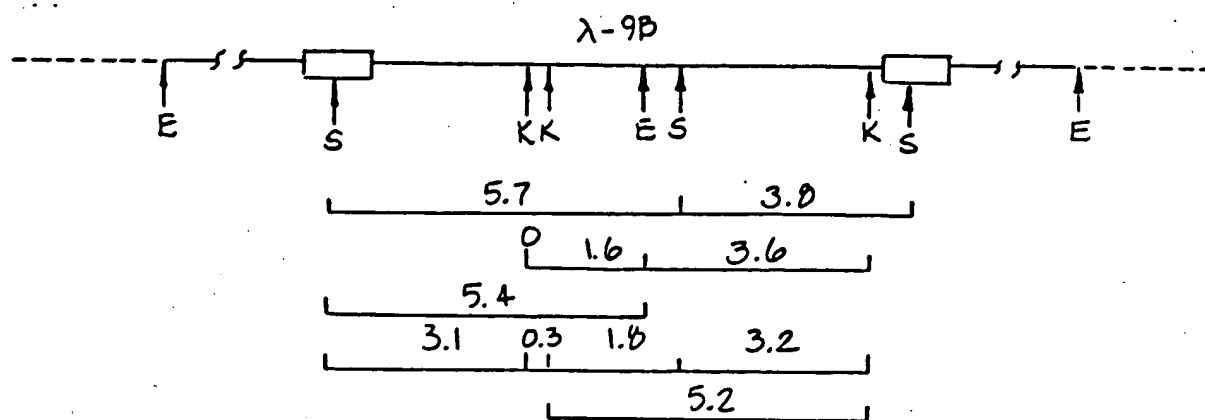


FIG. 2.

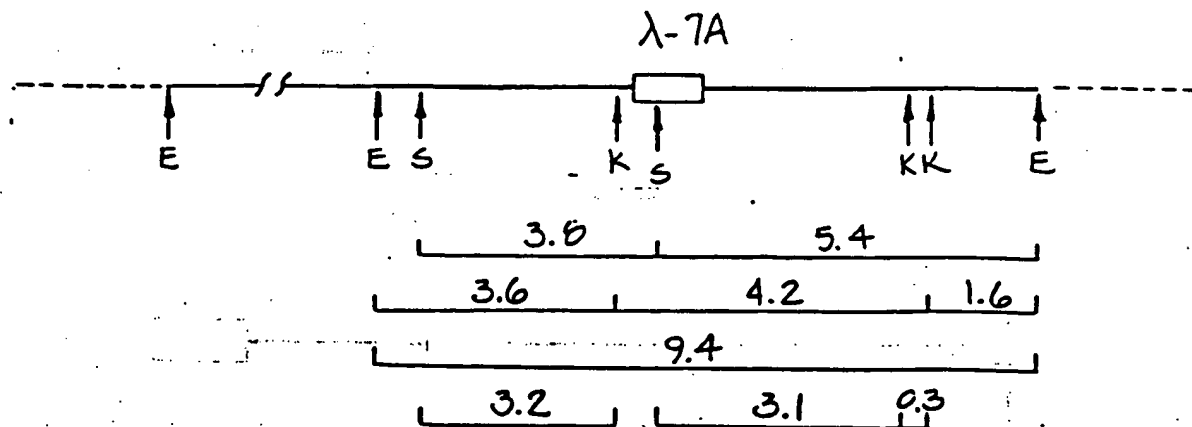
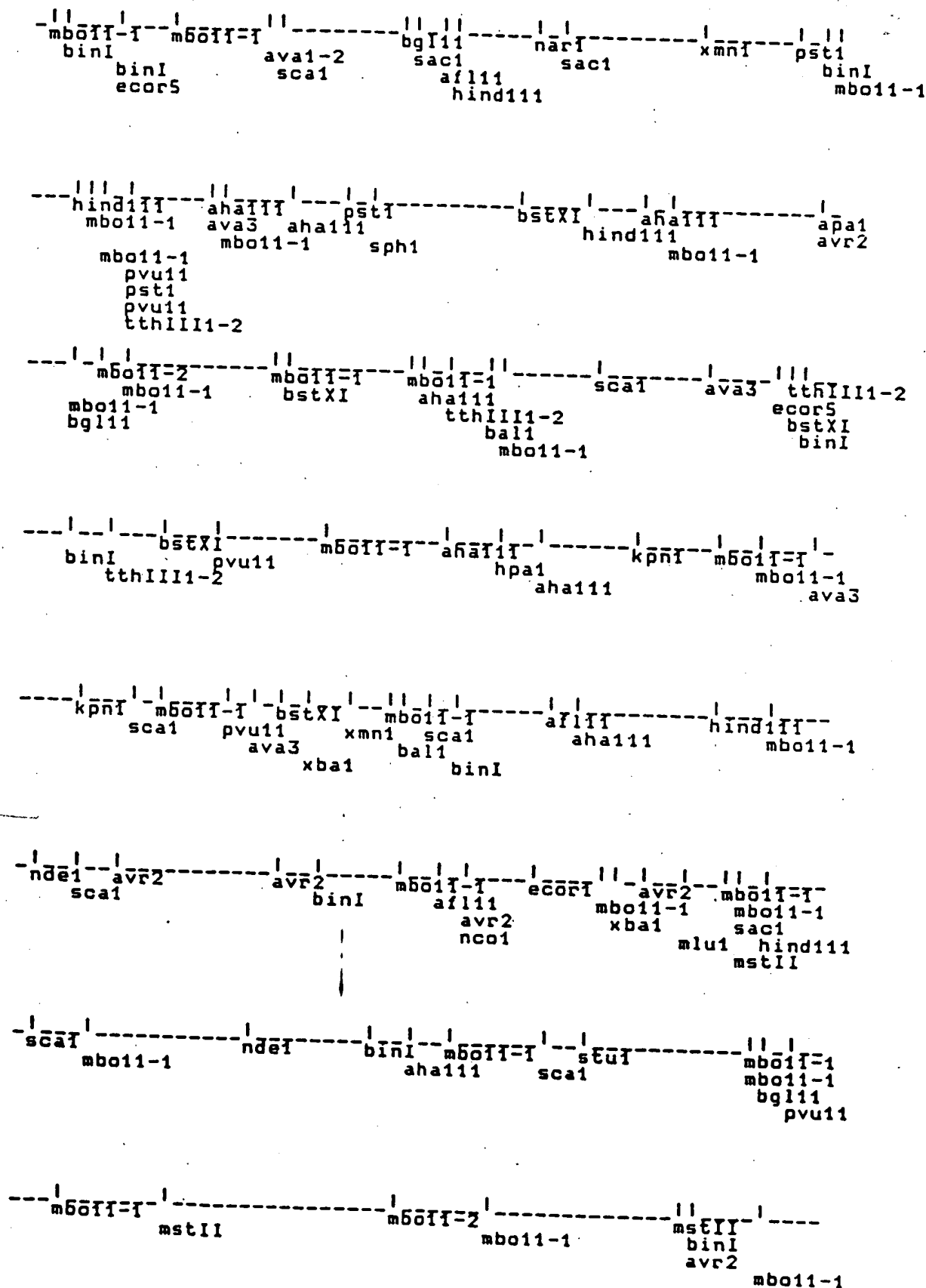


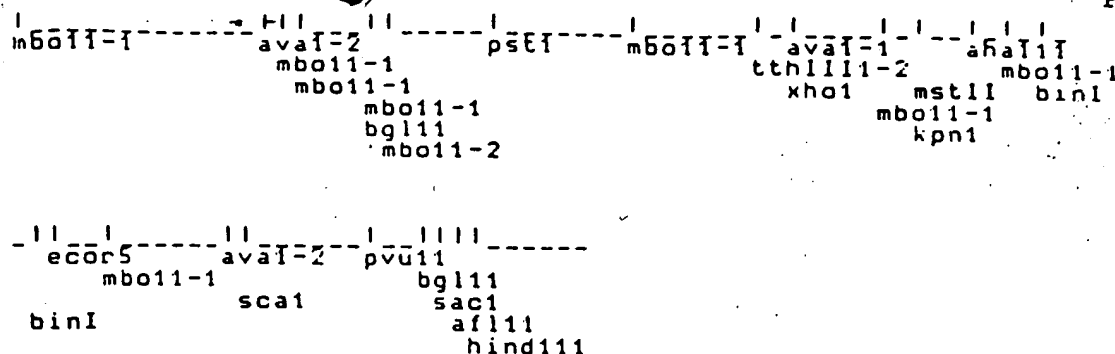
FIG. 3.



87 931154



931154



- 1 CTGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCACAC
GACCTTCCCATTAAACCAGGGTTTCTTCTGTTCTCTAGGAACCTAGACACCTAGATGGTGTG
26 mbo11, 50 bin1,
- 63 ACAAGGCTACTTCCCTGATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCACT
TGTTCCGATGAAGGGACTAACCCTCTTAATGTGTGGTCCCGGTCCCTAGTCTATAGGTGA
107 bin1, 113 ecor5,
- 123 GACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCAA
CTGGAAACCTACCACGAAGTTGATCATGGTCAACTCGGTCTCTCCATCTTCTCCGGTT
172 mbo11,
- 183 TGAAGGAGAGAAACAACAGCTTGTACACCCTATGAGCCTGCATGGGATGGAGGACGCGGA
ACTTCCTCTCTTGTGTGCAACAATGTGGGATACTCGGACGTACCCTACCTCCTGCGCCT
- 243 GAAAGAAGTGTAGTGTGGAGGTTTGACAGCAAACTAGCATTTTCATCACATGGCCCGAGA
CTTTCTTCACAATCACACCTCCAACTGTCTGTTGATCGTAAAGTAGTGTACCGGGCTCT
296 ava1,
- 303 GCTGCATCCGGAGTACTACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGC
CGACGTAGGCCTCATGATGTTTCTGACGACTGTAGCTCGAAAGATGTTCCCTGAAAGGCG
314 sca1,
- 363 TGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATGC
ACCCCTGAAAGGTCCCTCCGCACCGGACCCGCCCTGACCCCTCACCGCAGGGAGTCTACG
- 423 TGCATATAAGCAGACTGCTTTTTGCTGTACTGGGTCTCTCTGTTAGACCAGATCTGAG
ACGTATATTCTGCTGACGAAAAACGGACATGACCCAGAGAGACCAATCTGGTCTAGACTC
474 bgl11,
- 483 CCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTT
GGACCCTCGAGAGACCGATTGATCCCTTGGGTGACGAATTTCGGAGTTATTTGAAACGGAA.
488 sac1, 518 afl111, 532 hind111,
- 543 GAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCA
CTCACGAAGTTCATCACACACGGGCAGACAACACACTGAGACCATTGATCTCTAGGGAGT
- 603 GACCCTTTTAGTCACTGTGGAAAAATCTCTAGCAGTGGCGCCCGAACAGGGACGCGAAAG
CTGGGAAAAATCAGTCACACCTTTTATAGAGATCGTCACCGCGGGCTTGTCCCTGCGCTTTC
639 nar1,
- 663 CGAAAGTAGAACCAGAGGAAGCTCTCTCGACGCAAGGACTCGGCTTGCTGAAGCGCGCACAG
GCTTTCATCTTGGTCTCTCTCGAGAGAGCTGCGTCTGAGCCGAACGACTTCGCGCGTGTG
680 sac1,
- 723 CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAATTTTTGACTAGCGGAGGCTAGAAG
GTTCTCCGCTCCCCGCGCTGACCACTCATGCGGTAAAAACTGATCGCCTCCGATCTTC
- 783 GAGAGAGAGATGGGTGCGAGAGCGTCCGTATTAAGCGGGGAGAAATTAGATAAATGGGAA
CTCTCTCTACCCACGCTCTCGCAGCCATAATTGCCCCCTCTTAATCTATTTACCCTT

GAG

843 LysIleArgLeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAla
AAAATTTCGGTTAAAGGCCAGGGGGGAAAGAAAAATATAAGTTAAAACATATAGTATGGGCA
TTTTAAGCCAATTCCGGTCCCCCTTCTTTTTTATATTCAATTTTGTATATCATACCCGT

903 SerArgGluLeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCys
AGCAGGGAGCTAGAACGATTTCGCAGTCAATCCTGGCCTGTTAGAAACATCAGAAGGCTGC
TCGTCCCTCGATCTTGCTAAGCGTCAGTTAGGACCGACAATCTTTGTAGTCTTCCGACG

959 pst1,

963 ArgGlnIleLeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSer
AGACAAATATTGGGACAGCTACAGCCATCCCTTCAGACAGGATCAGAAGAACTTAGATCA
TCTGTTTATAACCTGTCTGATGTCTGGTAGGGAAGTCTGTCTAGTCTTCTTGAATCTAGT

1002 bin1, 1008 mbo11,

1023 LeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnArgIleAspValLysAspThr
TTATATAATACAGTAGCAACCCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACC
AATATATTATGTCATCGTTGGGAGATAACACATGTAGTTTCTATCTACATTTTCTGTGG

1083 LysGluAlaLeuGluLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGln
AAGGAAGCTTTAGAGAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAA
TTCCTTCGAAATCTCTTCTATCTCCTTCTCGTTTTGTTTTCTATTCTTTTTCCGTGTCTGT

1087 hind111, 1097 mbo11, 1107 mbo11, p25

1143 AlaAlaAlaAlaAlaGlyThrGlyAsnSerSerGlnValSerGlnAsnTyrProIleVal
GCAGCAGCTGCAGCTGGCACAGGAAACAGCAGCCAGGTCAGCCAAAATTACCCTATAGTG
CGTCGTCGACGTCGACCGTGTCTTTGTCTCGGTCCAGTCGGTTTTAATGGGATATCAC

1147 pvu11, 1150 pst1, 1153 pvu11, 1156 tthIII1,

1203 GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrp
CAGAACCTACAGGGGCAAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGG
GTCTTGGATGTCCCGTTTTACCATGTAGTCCGGTATAGTGGATCTTGAAATTTACGTACC

1250 aha111, 1255 ava3,

1263 ValLysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu
GTAAAAGTAGTAGAAGAAAAGGCTTTTCAGCCAGAAAGTAATACCCATGTTTTTCAGCATT
CATTTTCATCATCTTCTTTTCCGAAAGTCGGGTCTTCATTATGGGTACAAAAGTCGTAAT

1275 mbo11,

1323 SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGln
TCAGAAGGAGCCACCCACAAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAA
AGTCTTCTCTCGGTGGGGTGTCTAAATTTGTGGTACGATTTGTGTACCCCCCTGTAGTT

1346 aha111,

1383 AlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal
GCAGCCATGCAAAATGTTAAAGAGACTATCAATGAGGAAGCTGCAGAAATGGGATAGAGTG
CGTCGGTACGTTTACAATTTTCTCTGATAGTTACTCCTTCGACGTCTTACCCTATCTCAC

1423 pst1,

1443 HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAsp
CATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAC
GTAGGTCACGTACGTCCCGGATAACGTGGTCCGGTTTACTCTCTTGGTTCCCTTCACTG

1451 sph1,

1503 IleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro
ATAGCAGGAACCTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT
TATCGTCCTTGATGATCATGGGAAGTCCTTGTTTATCCTACCTACTGTTTATTAGGTGGA

1563 IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArg
ATCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGA
TAGGGTCATCCTCTTTAGATATTTTCTACCTATTAGGACCCTAATTTATTTTATCATTCT

1623 MetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp
ATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAGGAACCTTTAGAGAT
TACATATCGGGATGGTCGTAAGACCTGTATTCTGTTCTGGTTTCTTGGGAAATCTCTA

1636 bstXI,

1683 TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsn
TATGTAGACCGTTCTATAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAT
ATACATCTGGCCAAGATATTTTGAGATTCTCGGCTTGTTTGAAGTGTCTACATTTTTTA

1720 hind111,

931154

1743 TrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys
TGGATGACAGAAACCTTGTGGTCCAAATGCAAACCCAGATTGTAAGACTATTTTAAAA
ACCTACTGTCTTTGGAACAACCAGGTTTTACGTTTGGGTCTAACATTCTGATAAAATTTT
1796 aha111,

1803 AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGly
GCATTGGGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTGAGGGAGTGGGGGA
CGTAACCTGGTCTGTCGATGTGATCTTCTTTACTACTGTCGTACAGTCCCTCACCCCT
1827 mbo11,

1863 ProGlyHisLysAlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaAsn
CCCGGTCATAAAGCAAGAGTTTTGGCTGAAGCCATGAGCCAAGTAACAAATCCAGCTAAC
GGGCGGTTATTTCTCTCAAAACCGACTTCGGTACTCGGTTTCATTGTTTAGGTCGATTG
1923 IleMetMetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCys
ATAATGATGCAGAGAGGGCAATTTTAGGAACCAAAGAAAGACTGTTAAGTGTTCATTTGT
TATTACTACGTCTCTCCGTTAAATCCTTGGTTTCTTTCTGACAATTACAAAGTTAACA
1983 GlyLysGluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpArg
GGCAAAGAAGGGCACATAGCCAAAAATTCAGGGGCCCTAGGAAAAAGGGCTGTTGGAGA
CCGTTTCTTCCCGTGTATCGGTTTTTAACGTCCCGGGGATCCTTTTTCCCGACAACCTCT
2014 apa1, 2019 avr2,

2043 CysGlyArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGly
TGTGGAAGGGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGG
ACACCTTCCCTTCTGTGGTTTACTTTCTAACGTGACTCTCTGTCCGATTAAAAAATCCC
2102 mbo11,

2103 LysIleTrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluPro
AAGATCTGGCCTTCTCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGAGCCA
TTCTAGACCGGAAGGATGTTCCCTTCCGGTCCCTTAAAGAAGTCTCGTCTGGTCTCGGT
2104 bgl11, 2141 mbo11,

2163 ThrAlaProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLys
ACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAG
TGTGCGGGTGGTCTTCTCTCGAAGTCAAACCCCTCTCTTTTGTGAGGGAGAGTCTTC
2175 mbo11,

2223 GlnGluProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsn
CAGGAGCCGATAGACAAGGAAGTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAAC
GTCCTCGGCTATCTGTTCTTGACATAGGAAATTGAAGGGAGTCTAGTGAGAAACCGTTG
2283 AspProSerSerGlnOC
GACCCCTCGTCACAATAAGGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGA
CTGGGGAGCAGTGTATTCTATCCCCCGTTGATTTCTTCGAGATAATCTATGTCCT

2342 MetAsnLeuProGlyLysTrpLysProLysMetIle
GCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAATGGAAACCAAAATGATA
CGTCTACTATGTCATAATCTTCTTTACTTAAACGGTCTTTTACCTTTGGTTTTTACTAT
2360 mbo11, 2375 bstX1,

2402 GlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIleProValGluIleCys
GGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGT
CCCCCTTAACCTCCAAAATAGTTTCATTCTGTCTAGTCTATGGACATCTTTAGACA
2462 GlyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArg
GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGA
CCTGTATTTGATATCCATGTCTAATCATCCTGGATGTGGACAGTTGTATTAACCTTCT
2517 mbo11,

2522 AsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrVal
AATCTGTTGACTCAGATTGGTTGTACTTTAAATTTCCCATTAGTCTATTGAAACTGTA
TTAGACAACCTGAGTCTAACCAACATGAAATTTAAAGGGGTAATCAGGATAACTTTGACAT
2548 aha111, 2577 tth1111,

2582 ProValLysLeuLysProGlyMetAspGlyProLysValLysGlnTrpProLeuThrGlu
CCAGTAAAATTAAGCCAGGAATGGATGGCCCAAAGTTAAGCAATGGCCATTGACAGAA
GGTCATTTTAATTTCCGTCTTACCTACCGGGTTTTCAATTCGTTACCGGTAACCTGTCTT
2627 bal1, 2639 mbo11,

2642 GluLysIleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIleSer
GAAAAAATAAAGCATTAGTAGAGATATGTACAGAAATGGAAAAGGAAGGGAAAAATTTCA
CTTTTTTATTTTCGTAATCATCTCTATACATGCTTTTACCTTTTCTTCCCTTTTAAAGT

2702 LysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSer
AAAATTGGGCGCTGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAGACAGT
TTTTAACCCGGACTTTTAGGTATGTTATGAGGTCATAAACGATATTTCTTTTTCTGTCA
2759 sca1,
2762 ThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrp
ACTAAATGGAGAAAACCTAGATAGATTTAGAGAACTTAATAAAGAACTCAAGACTTCTGG
TGATTTACCTCTTTTGATCATCTAAAGTCTCTGAATTATTTCTTGAGTTCTGAAGACC
2822 GluValGlnLeuGlyIleProHisProGlnGlyOC
GAAGTTTCAGTTAGGAATACCAACCCGAGGGTTAAAAAAGAAAAAATCAGTAACAGTA
CTTCAAGTCAATCCTTATGGTGTGGGCGTCCCAATTTTTCTTTTTTAGTCATTGTCTCAT
2882 TTGGATGTGGGTGATGCATACCTTTTCAGTTCCTTAGATAAAGACTTTAGAAAGTATACTG
AACCTACACCCACTACGTATGAAAAGTCAAGGGAATCTATTTCTGAAATCTTTCATATGAC
2895 ava3,
2943 CATTTACCATACCTAGTATAAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGG
GTAAATGGTATGGATCATATTTGTTACTCTGTGGTCCCTAATCTATAGTCATGTTACACC
2985 ecor5,
3003 LeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeu
CTGCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTA
GACGGTGTCCCTACCTTTCTAGTGGTCTGTTATAAGGTTTCATCGTACTGTTTTTAGAAT
3003 tthIII1, 3006 bstXI, 3021 binI,
3063 GluProPheArgLysGlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyr
GAGCCTTTTAGAAAACAGAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTAT
CTCGGAAAATCTTTTGTCTTAGGTCTGTATCAATAGATAGTTATGTACCTACTAAACATA
3123 ValGlySerAspLeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHis
GTAGGATCTGACTTAGAAAATAGGGCAGCATAGAACAAAAATAGAGGAACTGAGACAGCAT
CATCCTAGACTGAATCTTTATCCCGTCTGATCTTGTTTTTATCTCCTTGACTCTGTCGTA
3126 binI, 3171 tthIII1,
3183 LeuLeuArgTrpGlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeu
CTGTTGAGGTGGGGATTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTT
GACAACTCCACCCCTAAATGGTGTGGTCTGTTTTTTGTAGTCTTCTTGAGGTAAGGAA
3234 bstXI,
3243 TrpMetGlyTyrGluLeuHisProAspLysTrpThrValGlnProIleMetLeuProGlu
TGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAATGCTGCCAGAA
ACCTACCCAATACTTGAGGTAGGACTATTTACCTGTCTGCGATATTACGACGGTCTT
3303 LysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSer
AAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAAATTGAATTGGGCAAGT
TTTCTGTCGACCTGACAGTTACTGTATGTCTTCAATCACCTTTTAACTTAACCCGTTCA
3308 pvu11,
3363 GlnIleTyrAlaGlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyThrLysAla
CAGATTTATGCAGGGATTAAAGTAAAGCAGTTATGTAAACTCCTTAGAGGAACCAAGCA
GTCTAAATACGTCCCTAATTTCAATTCGTCAATACATTTGAGGAATCTCCTTGGTTCTGT
3423 LeuThrGluValIleProLeuThrGluGluAlaGluLeuGluAlaGluAsnArgGlu
CTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAG
GATTGTCTTCATTATGGTGTATTGCTTCTTCTGCTCGATCTTGACCGTCTTTTGTCCCTC
3447 mbo11,
3483 IleLeuLysGluProValHisGluValTyrTyrAspProSerLysAspLeuValAlaGlu
ATTCTAAAAGAACAGTACATGAAGTATATTATGACCCATCAAAAGACTTAGTAGCAGAA
TAAGATTTTCTTGGTCATGTACTTCATATAATACTGGGTAGTTTTCTGAATCATCGTCTT
3543 IleGlnLysGlnGlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluProPheLysAsn
ATACAGAAGCAGGGGCAAGGGCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAAT
TATGTCTTCTGCCCCGTTCCGGTTACCTGTATAGTTTAAATAGTTCTCGGTAAATTTTAA
3594 aha111,
3603 LeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeu
CTGAAAACAGGAAAGTATGCAAGGATGAGGGGTGCCACACTAATGATGTAACACAGTTA
GACTTTTGTCTTTTCATACGTTCTTACTCCCCACGGGTGTGATTACTACATTTTGTCAAT
3659 hpa1,

POL

3663 ThrGluAlaValGlnLysValSerIleGluSerIleValIleTrpGlyLysIleProLys
ACAGAGGCAGTGCAGAAAAGTATCCACAGAAAGCATAGTAATATGGGGAAAGATTCTCTAAA
TGTCTCCGTACAGTTTTTTCATAGGTGCTTTTCGTATCATTATACCCCTTTCTAAGGATTT

3723 PheLysLeuProIleGlnLysGluThrTrpGluAlaTrpTrpMetGluTyrTrpGlnAla
TTTAAACTACCCATACAAAAGGAACATGGGAAGCATGGTGGATGGAGTATTGGCAAGCT
AAATTTGATGGGTATGTTTTCTTTGTACCCTTCGTACCACCTACCTCATAACCGTTCTGA
3723 aha111,

3783 ThrTrpIleProGluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGln
ACCTGGATTCTCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAATTATGGTACCAG
TGGACCTAAGGACTCACCCCTCAACAGTTATGGGGAGGGAATCACTTTAATACCATGGTC
3835 kpn1,

3843 LeuGluLysGluProIleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArg
TTAGAGAAAGAACCATAGTAGGAGCAGAACTTTCTATGTAGATGGGGCAGCTAATAGG
AATCTCTTTCTTGGGTATCATCCTCGTCTTTGAAAGATACATCTACCCCGTCGATTATCC

3903 GluThrLysLeuGlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSer
GAGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGTTGTCTCC
CTCTGATTTAATCCTTTTCGTCTATACAATGACTGTCTCTCTCTGTTTTTCAACAGAGG
3943 mbo11,

3963 IleAlaAspThrThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAsp
ATAGCTGACACAAACAAATCAGAAAGACTGAATTACAAGCAATTCATCTAGCTTTGCAGGAT
TATCGACTGTGTTGTTTAGTCTTCTGACTTAATGTTCTGTTAAGTAGATCGAAACGTCCTA
3983 mbo11,

4023 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAla
TCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCA
AGCCCTAATCTTCATTTGTATCATTGTCTGAGTGTATACGTAATCCTTAGTAAGTTCTGT
4060 ava3,

4083 GlnProAspLysSerGluSerGluLeuValSerGlnIleIleGluGlnLeuIleLysLys
CAACCAGATAAGAGTGAATCAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAG
GTTGGTCTATTCTCACTTAGTCTCAATCAGTCAGTTTATTATCTCGTCAATTATTTTTTC

4143 GluLysValTyrLeuAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnVal
GAAAAGGTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTA
CTTTTCCAGATGGACCGTACCCATGGTCGTGTGTTTCTTAACCTCCTTTACTTGTTCAT
4163 kpn1,

4203 AspLysLeuValSerAlaGlyIleArgLysValLeuPheLeuAsnGlyIleAspLysAla
GATAAATTAGTCAGTGTCTGGAATCAGGAAAGTACTATTTTGAATGGAATAGATAAGGCC
CTATTTAATCAGTCACGACCTTAGTCCTTTCATGATAAAAACTTACCTTATCTATTCCGG
4232 sca1,

4263 GlnGluGluHisGluLysTyrHisSerAsnTrpArgAlaMetAlaSerAspPheAsnLeu
CAAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTG
GTTCTTCTTGACTCTTTATAGTGTCTTAACCTCTCGTTACCGATCACTAAATTGGAC
4266 mbo11,

4323 ProProValValAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGlu
CCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGT CAGCTAAAAGGAGAA
GGTGGACATCATCGTTTTCTTTATCATCGGTCGACACTATTACAGTCGATTTTCTCTT
4352 pvu11,

4383 AlaMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeu
GCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATCTA
CGGTACGTACCTGTTTCATCTGCATCAGGTCCTTATACCGTTGATCTAACATGTGTAGAT
4386 ava3, 4410 bstXI, 4439 xba1,

4443 GluGlyLysIleIleLeuValAlaValHisValAlaSerGlyTyrIleGluAlaGluVal
GAAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTT
CTTCTTTTTTAATAGGACCATCGTCAAGTACATCGGTACCTATATATCTTCGTCTTCAA
4497 xmn1,

4503 IleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeuAlaGlyArgTrp
ATTCCAGCAGAGACAGGGCAGGAAACAGCATATTTTCTCTTAAATTAGCAGGAAGATGG
TAAGGTCTCTCTGTCCCGTCTTTGTCTGTATAAAAGAGAATTTTAATCGTCCTTCTACC
4555 mbo11, 4560 bal1,

4563 ProValLysThrIleHfsThrAspAsnGlySerAsnPheThrSerThrThrValLysAla
CCAGTAAAAACAATACATACAGACAATGGCAGCAATTTCCACAGTACTACGGTTAAGGCC
GGTCATTTTTTTGTTATGTATGTCTGTTACCGTCGTTAAAGTGGTCATGATGCCAATTCGGG
4605 sca1,
4623 AlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGln
GCCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAA
CGGACAACCCCGTCCCTAGTTCGTCTTAAACCGTAAGGGATGTTAGGGGTTTCAGTT
4639 bin1,
4683 GlyValValGluSerMetAsnAsnGluLeuLysLysIleIleGlyGlnValArgAspGln
GGAGTAGTAGAATCTATGAATAATGAATTAAGAAAATTATAGGACAGGTAAGAGATCAG
CCTCATCATCTTAGATACTTATTACTTAATTTCTTTTAATATCCTGTCCATTCTCTAGTC
4743 AlaGluHisLeuLysThrAlaValGlnMetAlaValPheIleHisAsnPheLysArgLys
GCTGAACACCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAGAAAA
CGACTTGTGGAATTCTGTCGTGATGTTTACCGTCATAAGTAGGTGTTAAATTTTCTTT
4752 af111, 4791 aha111,
4803 GlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleIleAlaThrAspIle
GGGGGGGATTGGGGGATACAGTGCAGGGGAAAGAAATAGTAGACATAATAGCAACAGACATA
CCCCCTAACCCCTATGTACGTCCTTTCTTATCATCTGTATTATCGTTGTCTGTAT
4863 GlnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArg
CAAACCTAAAGAACTACAAAAGCAAATTACAAAATTTCAAATTTTTCGGGTTTATTACAGG
GTTTGATTTCTTGATGTTTTCTGTTTAAATGTTTTTAAAGTTTAAAGCCCAAATAATGTCC
4923 AspAsnLysAspProLeuTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAla
GACAACAAAGATCCCTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAAGGGGCA
CTGTTGTTTCTAGGGGAAACCTTTCTGGTCTGTTTGAAGAGACCTTCCACTTCCCGT
4956 hind111,
4983 ValValIleGlnAspAsnSerAspIleLysValValProArgArgLysAlaLysIleIle
GTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAAGAAAGCAAAATCATT
CATCATTATGTTCTATTATCACTGTATTTTCATCACGGTTCTTCTTTTCGTTTTAGTAA
5023 mbo11,
5043 MetGluAsnArgTrpGlnValMetIleValTrpGlnValAspArgMetArgIle
ArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArgGlnAspGluAsp
AGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGAT
TCCCTAATACCTTTTGTCTACCGTCCACTACTAACACACCGTTCATCTGTCTACTCCTA
5103 ArgTrpTrpLysSerLeuValLysHisHisMetTyrIleSerLysLysAlaLysGlyTrp
AM
TAGAACATGGAAAAGTTTAGTAAACACCATATGTATATTTCAAAGAAAGCTAAAGGATGG
ATCTTGACCTTTTCAAATCATTTTGTGGTATACATATAAAGTTTCTTTCGATTTCTACC
5131 nde1,
5163 PheTyrArgHisHisTyrGluSerThrHisProArgValSerSerGluValHisIle
TTTTATAGACATCACTATGAAAGTACTCATCCAAGAGTAAGTTCAGAAGTACACATC
AAAATATCTGTAGTGATACTTTCATGAGTAGGTTCTCATTCAAGTCTTCATGTGTAG
5185 sca1,
5221 ProLeuGlyAspAlaLysLeuValIleThrThrTyrTrpGlyLeuHisThrGlyGluArg
CCCCTAGGGGATGCTAAATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGA
GGGGATCCCCTACGATTAAACCATTATTGTTGTATAACCCAGACGTATGTCCTCTTTCT
5223 avr2,
5281 GluTrpHisLeuGlyGlnGlyValAlaIleGluTrpArgLysLysLysTyrSerThrGln
GAATGGCATTTTGGGTCAGGGAGTCCGCATAGAATGGAGGAAAAAGAAATATAGCACACAA
CTACCGTAACCCGGTCCCTCAGCGGTATCTTACCTCCTTTTTCTTTATATCGTGTGTT
5341 ValAspProGlyLeuAlaAspGlnLeuIleHisLeuHisTyrPheAspCysPheSerGlu
GTAGACCCTGGCCTAGCAGACCAACTAATTCATCTGCATTATTTTGATTGTTTTTCAGAA
CATCTGGGACCGGATCGTCTGGTTGATTAAAGTAGACGTAATAAACTAACAAAAAGTCTT
5401 SerAlaIleLysAsnAlaIleLeuGlyTyrArgValSerProArgCysGluTyrGlnAla
TCTGCTATAAAAAATGCCATATTAGGATATAGAGTTAGTCCTAGGTGTGAATATCAAGCA
AGACGATATTTTTTACGGTATAATCCTATATCTCAATCAGGATCCACACTTATAGTTCTG
5440 avr2,
5461 GlyHisAsnLysValGlySerLeuGlnTyrLeuAlaLeuAlaAlaLeuIleThrProLys
GGACATAACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAAATAACACCAAAA
CCTGTATTGTTCCATCCTAGAGATGTTATGAACCGTGATCGTCTGAATTATTGTGGTTTT
5476 bin1,

5521 LysThrLysProProLeuProSerValLysLysLeuThrGluAspArgTrpAsnLysPro
AAGACAAAGCCACCTTTGCTAGTGTAAAGAACTGACAGAGGATAGATGGAACAAGCCC
TTCTGTTTCGGTGGAACGGATCACAATTCTTTGACTGTCTCCTATCTACCTTGTTCGGG

5581 GlnLysThrLysGlyHisArgGlySerHisThrMetAsnGlyHisAM
CAGAAGACCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGG
GTCTTCTGGTTCCCGGTGTCTCCCTCGGTATGTTACTTACCTGTGATCTCGAAAATCTCC

5583 mbo11,

5641 AGCTTAAGAGAGAAGCTGTTAGACATTTTCTAGGCCATGGCTCCATAGCTTAGGACAAT
TCGAATTCTCTCTTCGACAATCTGTAAAAGGATCCGGTACCGAGGTATCGAATCCTGTTA

5643 afl111, 5670 avr2, 5676 nco1,

5701 ATATCTATGAACTTATGGGGATACTTGGGCAGGAGTGGGAAGCCATAATAAGAATTCTGC
TATAGATACTTTGAATACCCCTATGAACCCGTCCTCACCTTCGGTATTATTCTTAAGACG

5752 ecor1,

5761 AACAACTGCTGTTTATTCATTTTCAAGATTGGGTGTCAACATAGCAGAATAGGCATTATTC
TTGTTGACGACAAATAAGTAAAGTCTTAACCCACAGTTGTATCGTCTTATCCGTAATAAG

5821 AACAGAGGAGAGCAAGAAGAAATGGAGCCAGTAGATCCTAATCTAGAGCCCTGGAAGCAT
TTGTCTCTCTCGTTCTTTTACCTCGGTATCTAGGATTAGATCTCGGGACCTTCGTA

5836 mbo11, 5862 xba1,

5881 CCAGGAAGTCAGCCTAGGACTGCTTGTAAACAATTGCTATTGTAAAAAGTGTGCTTTTCAT
GGTCCTTCAGTCGGATCCTGACGAACATTGTTAACGATAACATTTTTCAACGAAAGTA

5893 avr2,

5941 TGCTACGCGTGTTCACAAGAAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGA
ACGATGCGCACAAAGTGTCTTTTCCGAATCCGTAGAGGATACCGTCTTCTTCGCCTCT

5945 mlu1, 5988 mbo11,

6001 CAGCGACGAAGAGCTCCTCAGGACAGTCAGACTCATCAAGCTTCTCTATCAAGCAGTAA
GTCGCTGCTTCTCGAGGAGTCTGTGAGTCTGAGTAGTTCGAAGAGATAGTTTCGTCATT

6008 mbo11, 6011 sac1, 6016 mstII, 6038 hind111,

6061 GTAGTAAATGTAATGCAATCTTTACAAATATTAGCAATAGTATCATTAGTAGTAGTAGCA
CATCATTTACATTACGTTAGAAATGTTTATAATCGTTATCATAGTAATCATCATCATCGT

6121 ATAATAGCAATAGTTGTGTGGACCATAGTACTCATAGAATATAGGAAAATATTAAGACAA
TATTATCGTTATCAACACACCTGGTATCATGAGTATCTTATATCCTTTTATAATTCTGTT

6147 sca1,

6181 AGAAAATAGACAGATTAATTGATAGAATAAGAGAAAAAGCAGAAGACAGTGGCAATGAAA MetLys
TCTTTTATCTGTCTAATTAACATCTTATTCTCTTTTTCGTTCTTCTGTACCCGTTACTTT ENV

6222 mbo11,

6241 ValLysGlyThrArgArgAsnTyrGlnHisLeuTrpArgTrpGlyThrLeuLeuLeuGly
GTGAAGGGGACCAGGAGGAATTATCAGCACTTGTGGAGATGGGGCACCTTGCTCCTTGGG
CACTTCCCCTGGTCTCTTAATAGTCGTGAACACCTCTACCCCGTGGAACGAGGAACCC

6301 MetLeuMetIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValPro
ATGTTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTCACAGTTTATTATGGAGTACCT
TACAACACTAGACATCAGGATGTCTTTTAAACCCAGTGTCAAATAATACCTCATGGA

6361 ValTrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAlaArgAlaTyrAspThr
GTGTGGAAAGAAGCAACTACCACTCTATTTTGTGCATCAGATGCTAGAGCATATGATACA
CACACCTTTCTTCTTGATGGTGAGATAAAACACGTAGTCTACGATCTCGTATACTATGT

6410 nde1,

6421 GluValHisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGlu
GAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAA
CTCCATGTATTACAAACCCGGTGTGTACGGACACATGGGTGTCTGGGGTTGGGTGTTCTT

6481 ValValLeuGlyAsnValThrGluAspPheAsnMetTrpLysAsnAsnMetValGluGln
GTAGTATTGGGAAATGTGACAGAAAAATTTTAAACATGTGGAAAAATAACATGGTAGAACAG
CATCATAACCCCTTTACACTGTCTTTTAAAAATTGTACACCTTTTATTGTACCATCTTGTC

6541 MetGlnGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysValLysLeuThr
ATGCAGGAGGATATAATCAGTTTATGGGATCAAGCCTAAAGCCATGTGTAAAATTAACC
TACGTCCTCTATATTAGTCAAATACCCTAGTTTTCGGATTTCGGTACACATTTTAATTGG

6567 bni1,

5601 ProLeuCysValThrLeuAsnCysThrAspLeuGlyLysAlaThrAsnThrAsnSerSer
CCACTCTGTGTTACTTTAAATTGCACTGATTTGGGGAAGGCTACTAATACCAATAGTAGT
GGTGAGACACAATGAAATTTAACGTGACTAAACCCCTTCCGATGATTATGGTTATCATCA
6615 aha111,
5661 AsnTrpLysGluGluIleLysGlyGluIleLysAsnCysSerPheAsnIleThrThrSer
AATTGGAAAGAAGAAATAAAAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGC
TTAACCTTTCTTTATTTTCTCTTTATTTTGTGACGAGAAAGTTATAGTGGTGTTCG
6670 mbo11,
5721 IleArgAspLysIleGlnLysGluAsnAlaLeuPheArgAsnLeuAspValValProIle
ATAAGAGATAAGATTGAGAAAGAAAATGCACTTTTTTCGTAACTTTGATGTAGTACCAATA
TATTCTCTATTCTAAGTCTTTCTTTTACGTGAAAAAGCATTGGAACATCATGGTTAT
5781 AspAsnAlaSerThrThrThrAsnTyrThrAsnTyrArgLeuIleHisCysAsnArgSer
GATAATGCTAGTACTACTACCAACTATACCAACTATAGGTTGATACATTGTAACAGATCA
CTATTACGATCATGATGATGGTTGATATGGTTGATATCCAACATGTAACATTGTCTAGT
6790 sca1,
5841 ValIleThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCysThr
GTCATTACACAGGCCTGTCCAAAGGTATCATTGTAGCCAAATCCCATACATTATTGTACC
CAGTAATGTGTCGGACAGGTTTCCATAGTAAACTCGGTTAAGGGTATGTAATAACATGG
6851 stu1,
5901 ProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyLysGlyProCys
CCGGCTGGTTTTGCGATTCTAAAGTGTAATAATAAAACGTTCAATGGAAAAGGACCATGT
GGCCGACCAAAACGCTAAGATTTACATTATTATTTTGCAAGTTACCTTTTCTGGTACA
5961 ThrAsnValSerThrValGlnCysThrHisGlyIleArgProIleValSerThrGlnLeu
ACAAATGTGAGCACAGTACAATGTACACATGGAATTAGGCCAATAGTGTCAACTCAACTG
TGTTACAGTCGTGTCATGTTACATGTGTACCTTAATCCGGTTATCACAGTTGAGTTGAC
7021 LeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerAspAsnPheThrAsn
CTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTACGAAC
GACAATTTACCGTCAGATCGTCTTCTCTCCATCATTAACTAGACTGTTAAAGTGCTTG
7042 mbo11, 7045 mbo11, 7060 bgl11,
7081 AsnAlaLysThrIleIleValGlnLeuAsnGluSerValAlaIleAsnCysThrArgPro
AATGCTAAAACCATAATAGTACAGCTGAATGAATCTGTAGCAATTAAGTGTACAAGACCC
TTACGATTTTGGTATTATCATGTCGACTTACTTAGACATCGTTAATTGACATGTTCTGGG
7102 pvu11,
7141 AsnAsnAsnThrArgLysSerIleTyrIleGlyProGlyArgAlaPheHisThrThrGly
AACACAATACAAGAAAAAGTATCTATATAGGACCAGGGAGAGCATTTCATACAACAGGA
TTGTTGTTATGTTCTTTTTTCATAGATATATCCTGGTCCCTCTCGTAAAGTATGTTGTCCT
7199 mbo11,
7201 ArgIleIleGlyAspIleArgLysAlaHisCysAsnIleSerArgAlaGlnTrpAsnAsn
AGAATAATAGGAGATATAAGAAAAGCACATTGTAACATTAGTAGAGCACAAATGGAATAAC
TCTTATTATCCTCTATATTCTTTTCTGTGTAACATTGTAATCATCTCGTGTACCTTATTG
7261 ThrLeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsnAsnLysThrIleVal
ACTTTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGGAATAATAAACAAATAGTC
TGAAATCTTGCTATCAATTTTTTAATTCTCTTGTCAAACCCCTTATTATTTGTTATCAG
7321 PheAsnGlnSerSerGlyGlyAspProGluIleValMetHisSerPheAsnCysArgGly
TTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTAAATTGTAGAGGG
AAATTAGTTAGGAGTCCTCCCTGGGTCTTTAACATTACGTGTCAAATTAACATCTCCC
7331 mstII,
7381 GluPhePheTyrCysAsnThrThrGlnLeuPheAsnAsnThrTrpArgLeuAsnHisThr
GAATTTTTCTACTGTAATACAACACAAGTGTAAATAATACATGGAGGTTAAATCACACT
CTTAAAAAGATGACATTATGTTGTGTTGACAAATTATTATGTACCTCCAATTTAGTGTGA
7441 GluGlyThrLysGlyAsnAspThrIleIleLeuProCysArgIleLysGlnIleIleAsn
GAAGGAACATAAGGAATGACACAATCATACTCCCATGTAGAATAAAACAAATTATAAAC
CTTCCTTGATTCTCTTACTGTGTAGTATGAGGGTACATCTTATTTTGTTTAATATTG
7501 MetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGlyGlyGlnIleSerCys
ATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATTTGGAGGACAAATTAGTTGT
TACACCGTCCTTCATCCTTTTCTGTACATACGGGGAGGGTAACCTCCTGTTTAAATCAACA
7561 SerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThrAsnValThrAsnAsp
TCATCAAATATTACAGGGCTGCTATTAAACAAGAGATGGTGGTACAAATGTAACATAATGAC
AGTAGTTTATAATGTCCCGACGATAATTGTTCTCTACCACCATGTTTACATTGATTACTG

7621 ThrGluValPheArgProGlyGlyGlyAspMetArgAspAsnTrpArgSerGluLeuTyr
ACCGAGGTCTTCAGACCTGGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATAT
TGGCTCCAGAAGTCTGGACCTCCTCCTCTATACTCCCTGTTAACCTCTTCACTTAATATA
7628 mbo11,
7681 LysTyrLysValIleLysIleGluProLeuGlyIleAlaProThrLysAlaLysArgArg
AAATATAAAGTAATAAAATTGAACCATAGGAATAGCACCCACCAAGGCAAAGAGAGA
TTTATATTTTATTATTTTAACTTGGTAATCCTTATCGTGGGTGGTTCCGTTTCTCTTCT
7736 mbo11,
7741 ValValGlnArgGluLysArgAlaValGlyIleValGlyAlaMetPheLeuGlyPheLeu
GTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGTAGGAGCTATGTTCTTGGGTCTTGT
CACCACGTCTCTCTTTTTTCTCGTCACCCCTTATCATCCTCGATACAAGGAACCCAAGAAC
7801 GlyAlaAlaGlySerThrMetGlyAlaValSerLeuThrLeuThrValGlnAlaArgGln
GGAGCAGCAGGAAGCACTATGGGCGCAGTGTCTTACGCTGACGGTACAGGCCAGACAA
CCTCGTCGTCTTCTGTGATACCCGCGTCACAGTAACTGCGACTGCCATGTCCGGTCTGTT
7861 LeuLeuSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGln
TTATTGTCTGGTATAGTGCACAGCAGAAACAATTGCTGAGGGCTATTGAGGCGCAACAA
AATAACAGACCATATCAGTTGTCTGTTGTTAAACGACTCCCGATAACTCCGCGTTGTT
7921 HisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaVal
CATCTGTTGCAACTCACAGTCTGGGCGCATCAAGCAGCTCCAGGCAAGAGTCTGGCTGTG
GTAGACAACGTTGAGTGTGACACCCCGTAGTTCTGTCGAGGTCCGTTCTCAGGACCGACAC
7981 GluArgTyrLeuArgAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIle
GAAAGATACCTAAGGGATCAACAGCTCCTAGGGATTGGGGTTGCTCTGGAAACTCATT
CTTTCTATGGATTCCCTAGTTGTGCGAGGATCCCTAAACCCCAACGAGACCTTTTGAGTAA
7989 mstII, 7995 binI, 8007 avr2,
8041 CysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluAspIleTrp
TGCACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAAGACATTTGG
ACGTGGTGACGACACGGAACCTTACGATCAACCTCATTATTTAGAGACCTTCTGTAAACC
8089 mbo11,
8101 AspAsnMetThrTrpMetGlnTrpGluArgGluIleAspAsnTyrThrAsnThrIleTyr
GATAACATGACCTGGATGCAGTGGGAAAGAGAAATTGACAATTACACAAACACAATATAC
CTATTGTACTGGACCTACGTACCCTTTCTTTAACTGTTAATGTGTTTGTGTTATATG
8161 ThrLeuLeuGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu
ACCTTACTTGAAGAATCGCAGAACCAACAAGAAAGAAATGAACAAGAATTATTAGAATTG
TGAATGAACCTTCTAGCGTCTTGGTTGTTCTTTCTTACTTGTCTTAATAATCTTAAC
8170 mbo11,
8221 AspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrAsnTrpLeuTrpTyrIleLys
GATAAGTGGGCAAGTTTGTGGAATTGTTTAGCATAACAACTGGCTGTGGTATATAAAG
CTATTCACCCGTTCAAACACCTTAACCAAATCGTATTGTTGACCGACACCATATATTC
8281 IlePheIleMetIleValGlyGlyLeuValGlyLeuArgIleValPheAlaValLeuSer
ATATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGTGCTTTCT
TATAAGTATTACTATCATCTCCGAACCATCAAATTCTTATCAAAAACGACAGAAAGA
8341 IleValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrArgLeuProVal
ATAGTGAATAGAGTTAGGCAGGGATACTCACCATTGTCTTTTCTGAGACCGCCTCCAGTC
TATCACTTATCTCAATCCGTCCCTATGAGTGGTAACAGTAAAGTCTGGGCGGAGGGTCAG
8400 avai,
8401 ProArgGlyProAspArgProAspGlyIleGluGluGluGlyGlyGluArgAspArgAsp
CCGAGGGGACCCGACAGGCCCGACGGAATCGAAGAAGAAGGTGGAGAGAGAGACAGAGAC
GGCTCCCTGGGCTGTCCGGGCTGCCCTTAGCTTCTTCTTCCACCTCTCTCTGTCTGT
8431 mbo11, 8434 mbo11,
8461 ArgSerValArgLeuValAspGlyPheLeuAlaLeuIleTrpGluAspLeuArgSerLeu
AGATCCGTTCTGATTAGTGGATGGATTCTTAGCACTTATCTGGGAAGATCTGCGGAGCCTG
TCTAGGCAAGCTAATCACCTACCTAAGAATCGTGAATAGACCTTCTAGACGCCTCGGAC
8503 mbo11, 8505 bg111,
8521 CysLeuPheSerTyrArgArgLeuArgAspLeuLeuLeuIleAlaAlaArgThrValGlu
TGCTCTTTCAGCTACCGCCGCTTGGAGAGACTTACTCTTGATTGACGCGAGGACTGTGGAA
ACGGAGAAGTCTGATGGCGCGAACTCTCTGAATGAGAACTAACGTGCTCCTGACACCTT
8525 mbo11,

3581 IleLeuGlyHisArgGlyTrpG...AlaLeuLysTyrTrpTrpSerLeuLeuG...nTyrTrp
ATTCTGGGGCACAGGGGGTGGGAAGCCCTCAAATATTGGTGGAGTCTCCTGCAGTATTGG
TAAGACCCCGTGTCCCCACCCCTTCGGGAGTTTATAACCACCTCAGAGGACGTCAATAACC
8629 pst1,
3641 IleGlnGluLeuLysAsnSerAlaValSerTrpLeuAsnAlaThrAlaIleAlaValThr
ATTCAGGAACATAAAGAATAGTGCTGTAGCTGGCTCAACGCCACAGCTATAGCAGTAACT
TAAGTCCTTGATTCTTATCACGACAATCGACCGAGTTGCGGTGTGATATCGTCATTGA
3701 GluGlyThrAspArgValIleGluValAlaGlnArgAlaTyrArgAlaIleLeuHisIle
GAGGGGACAGATAGGGTTATAGAAGTAGCACAAAGAGCTTATAGAGCTATTCTCCACATA
CTCCCTGTCTATCCCAATATCTTCATCGTGTCTCGAATATCTCGATAAGAGGTGTAT
3761 HisArgArgIleArgGlnGlyLeuGluArgLeuLeuLeuOC MetGlyGlyLysTrpSer
CATAGAAGAATTAGACAGGGCTTGGAAAGGCTTTTGCTATAAGATGGGTGGCAAGTGGTCA
GTATCTTCTTAATCTGTCCCGAACCTTTCCGAAAACGATATTCTACCCACCGTTCACCACT
8765 mbo11,
3822 LysArgSerMetGlyGlyTrpSerAlaIleArgGluArgMetArgArgAlaGluProArg
AAACGTAGTATGGGTGGATGGTCTGCTATAAGGGAAAGAAATGAGACGAGCTGAGCCACGA
TTTGCATCATACCCACCTACCAGACGATATTCCCTTTCTACTCTGCTCGACTCGGTGCT
3882 AlaGluProAlaAlaAspGlyValGlyAlaValSerArgAspLeuGluLysHisGlyAla
GCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTGGAAAAACATGGAGCA
CGACTCGGTGTCGTCTACCCACCCCTCGTCATAGAGCTCTGGACCTTTTGTACCTCGT
8883 tthIII1, 8916 ava1 xho1,
3942 IleThrSerSerAsnThrAlaAlaThrAsnAlaAspCysAlaTrpLeuGluAlaGlnGlu
ATCACAAGTAGCAATACAGCAGCTACTAATGCTGATTGTGCTGGCTAGAAGCACAAAGAG
TAGTGTTTCATCGTTATGTCTGTCGATGATTACGACTAACACGGACCGATCTTCGTGTTCTC
9002 GluGluGluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLys
GAGGAAGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGACTTACAAG
CTCCTTCTCCACCCAAAAGGTGAGTCTGGAGTCCATGGAAATTCTGGTACTGAATGTTG
9005 mbo11, 9029 mstII, 9034 kpn1,
9062 AlaAlaLeuAspIleSerHisPheLeuLysGluLysGlyGlyLeuGluGlyLeuIleTrp
GCAGCTTTAGATATTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTTGG
CGTCGAAATCTATAATCGGTGAAAAATTTCTTTTCCCCCTGACCTTCCCGATTAAACC
9085 aha111,
9122 SerGlnArgArgGlnGluIleLeuAspLeuTrpIleTyrHisThrGlnGlyTyrPhePro
TCCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCT
AGGGTTTCTTCTGTTCTCTAGGAAGTAGACACCTAGATGGTGTGTGTTCCGATGAAGGGA
9129 mbo11, 9153 binI,
9182 AspTrpGlnAsnTyrThrProGlyProGlyIleArgTyrProLeuThrPheGlyTrpCys
GATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGC
CTAACCGTCTTAATGTGTGGTCCCGGTCCCTAGTCTATAGGTGACTGGAAACCTACCAG
9210 binI, 9216 ecor5,
9242 PheLysLeuValProValGluProGluLysValGluGluAlaAsnGluGlyGluAsnAsn
TTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCAATGAAGGAGAGAAACAAC
AAGTTCGATCATGGTCAACTCGGTCTCTTCATCTTCTCCGGTTACTTCTCTCTTGTG
9275 mbo11,
9302 SerLeuLeuHisProMetSerLeuHisGlyMetGluAspAlaGluLysGluValLeuVal
AGCTTGTACACCTATGAGCCTGCATGGGATGGAGGACGCGGAGAAAGAAGTGTAGTG
TCGAACAATGTGGGATACTCGGACGTACCCTACCTCTGCGCCTCTTTCTTCACAATCAC
9362 TrpArgPheAspSerLysLeuAlaPheHisHisMetAlaArgGluLeuHisProGluTyr
TGGAGGTTTGACAGCAAACCTAGCATTTTCATCAGTGGCCCGAGAGCTGCATCCGGAGTAC
ACCTCCAAACTGTCTGTTTGTATCGTAAAGTAGTGTACCGGGCTCTCGACGTAGGCCCTCATG
9399 ava1, 9417 sca1,
9422 TyrLysAspCysOP
TACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGG
ATGTTTCTGACGACTGTAGCTCGAAAGATGTTCCCTGAAAGGCGACCCCTGAAAGGTCCC
9482 AGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATGCTGCATATAAGCAGCTG
TCCGCACCGGACCCGCTGACCCCTCACCAGGGAGTCTACGACGTATATTCTGTCGAC
9536 pvu11,

9542 CTTTTTGCCTGTACTGGGTCTC...GGTTAGACCAGATCTGAGCCTGGGAGC...TCTGGC
GAAAAACGGACATGACCCAGAGAGACCAATCTGGTCTAGACTCGGACCCTCGAGAGACCG
9576 bg111, 9590 sac1,
9602 TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTG
ATTGATCCCTTGGGTGACGAATTCGGAGTTATTCGAACGGAACCTCACGAAGTTCATCAC
9620 af111, 9634 hind111,
9662 TGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTG
ACACGGGCAGACAACACACTGAGACCATTGATCTCTAGGGAGTCTGGGAAAATCAGTCAC
9722 TGGAAAAATCTCTAGCAG
ACCTTTTtagagatcgTC

15

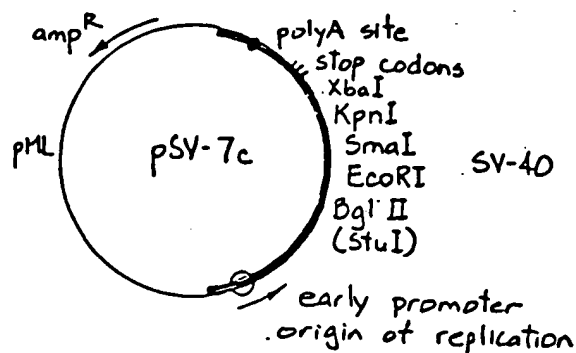
FIGURE 5

FIGURE 5
3 of 4

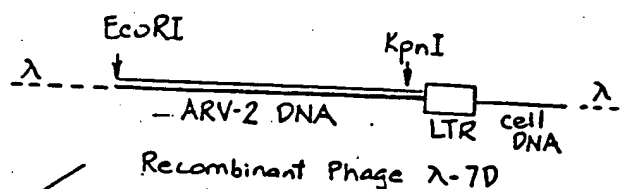
۱۲۲

77708 LeuLeuGluGluSerGlnAaenGlnGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrpPheserIleThrAsnTrpLeuTrpTyrlleLysIle 683
 TTTACTTGGAAGANTCGCAGAACCAACAGAAAGANTGAACAAGAAATATTAGAAATTTGGTGGCAAGTTGTGGAAATTTGTAGCATAAACAACACTGGCTGTGGTATATTAAGANTA
 7828 pheIleMetIleValGlyLeuValGlyLeuArgIleValIlePheAlaValLeuSerIleValAsnArgValArgGlnGlyTrpSerProLeuSerPheGlnThrArgLeuProValPro 723
 TTCATAATGATAGTAGAGGGCTTGGTAGGTTTAAAGATAGTTTTCGTGTGCTTTCATTAGTAGATAGAGTTAGGCAGGATCTCACACCATTTGCTCATTTTCAGCCGCCCTCCCGAGTCCCG
 7948 ArgGlyProAspArgProAspGlyIleGluGluGluGlyGlyGluArgAspArgAspArgValArgLeuValAspGlyLeuAlaLeuLeuTrpGluAspLeuArgSerLeuCys 763
 AGGGGCCCGCAGCAGCCCGCGGAAATCCGAGAGAGAGAGAGACAGACAGATCCGTTCCATTAGTAGTGGAATTCATTAGGCATTTATCTGGGAAGATCTGCGGAGCCTGTGCG
 8068 LeuPheSerTyArgArgLeuArgAspLeuLeuAlaAlaArgThrValGluIleLeuGlyHisArgGlyTrpGluAlaLeuLysTyTrpTrpSerLeuLeuGlnTyTrpIle 803
 CTCCTTCAGCTACCCGCCCTTCAGAGACTTACTTCATTGTCAGCGAGGACTGTGGAAATTCCTGGGGCACAGGGGGTGGGAAGCCCTCAATANTTGTGGAGTCTCTCCGACGATATTGGANT
 8188 GlnGluLeuLysAsnSerAlaValSerTrpLeuAsnAlaThrAlaIleAlaValThrGluGlyThrAspArgValIleGluValAlaGlnArgAlaTyArgAlaIleLeuHisIleHis 843
 CAGGAACATAAGAATAGTGTCTTATGCTGGCTCAACGCCACGCTATAGCAGTAACTGAGGGGACAGATAGGGTTATAGAGTAGCACAAAGAGCTTTATAGAGCTATTCTCCACATACAT
 ArgArgIleArgGlnGlyLeuLeuLeuLeuOC
 8308 AGAAGATTTAGACAGGGCTTGGAAAGGCTTTTGCTATTAAAGATGGGTGGCAAGTGGTTCAAAACGTAGTATGGTGTGATGGTCTCTATTATAGGGAAGAAGTACAGAGCTGAGCCACGAGC
 8428 TGAGCCAGCAGCAGANTGGGTGGGAGCAGTATCTCGAGAGCCTGGAAAAACATGAGCAANTCACAAGTAGCAATACAGCAGCTACTAATGCTGATTTGTCCCTGGCTAGAGAAGCACAGAGGA
 8548 GGAAGAGGTGGGTTTTCAGATCAGAGACCTCAGGTACCTTTTAAGACCATGACTTACAGGCAGCTTTAGTATATTAGCCACTTTTAAAGAAAGGGGGGA CTGGAAGGGCTTAATTTGCT
 8667 CCCAAAGAGACAAAGAGATCCTTGATCTGTGGATGTACCACACACAGGCTACTTCCCTGATTTGCCAAGATTTACACACAGGCCACGGGATCAGATATCCACCTGACCTTTTGGATGGTCT
 8787 TCAGCTAGTACAGTTCAGCCAGAGAGAGAGAGGCCAATGAAGGAGAGACACACAGCTTGTGTACACCTATAGCCCTCATGGGATGGAGGACGGGAGAAAGAAAGTGTTAGTGT
 8907 GGAGGTTTGCACGACAACTAGCANTTTCATCATCATGTCGCCGAGAGCTGCATCCGGGATCTACAAAGACTGCTGCACATCGAGCTTTCTACAAAGGAGACTTTTCGGCTGGGGAGCTTTTCACAGGGA
 9027 GGCCTGGCCTGGCGGGAGTGGGGAGTGGCGTCCCTCAGATGCTGCATATTAAGCAGCTGCTTTTTCGCTGTACTG GGTCTCTCTGTGTAGACCAAGATCTGAGCCTGGGAGCTCTCTGGC
 9146 TTACTTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCTTGGAGTGCTTCA AGTAGTGTGGCCGCTCTGTTGTGTACTCTGTGTAACTAGAGATCCCTCAGACCCCTTTTATGTCAGT
 9265 GTGGAAAAAATCTTAGCAG
 ← U5
 ← U3 R →
 ← R U5 →

FIGURE 5
4 of 4



digestion with
KpnI and EcoRI



digestion with
EcoRI and KpnI

ligation

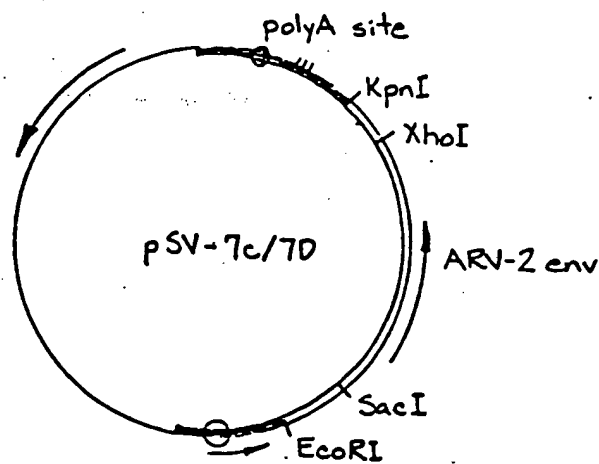


FIG. 6.

Figure 7

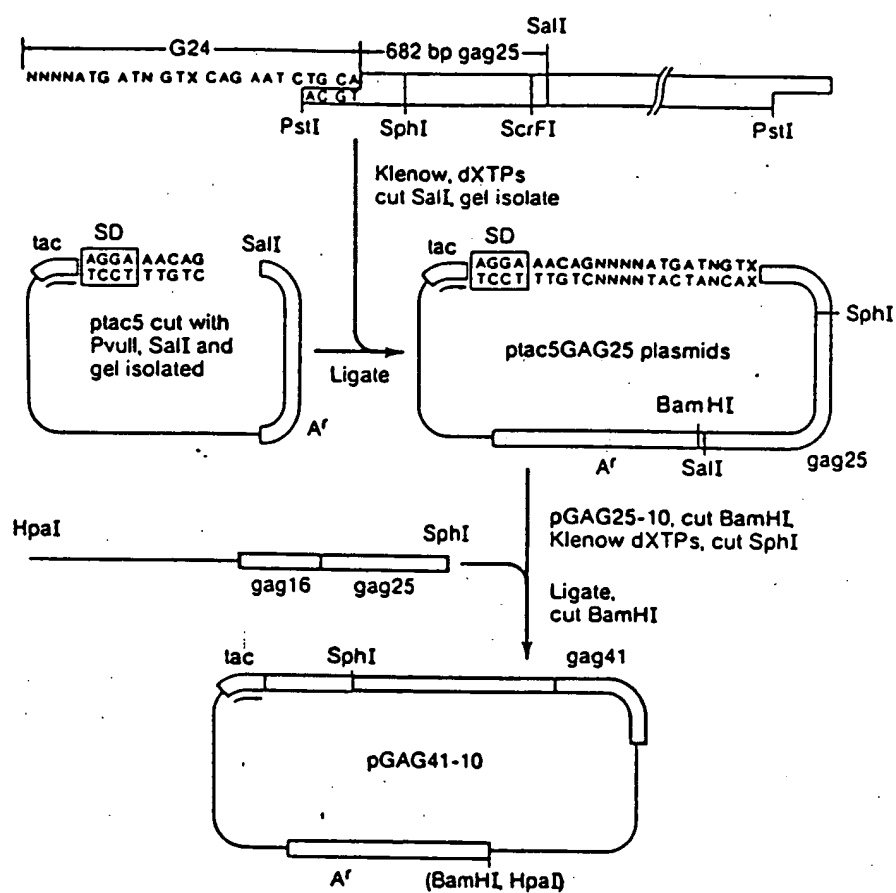


FIG. 8

	ptac 5 Promotor	* Met I/2 Val ATG ATC GTA
748	GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu	181
	CAGAATCTGCAGGGGCCAAATGGTACATCAGGCCCATATCACCCTAGAACCTTTAAATGCATCGCTAAAGTAGTAGAAGAAAGGCCTTCACCCACAGAACTAATACCCATGTTTTCAGCATT	
868	SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal	221
	TCAGAAGGAGCCACCCACAAAGATTTAAACACCATCTCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAATGTTTAAAGAGACTATCAATGAGGAAGCTCCAGAAATGGGATAGAGTG	
988	HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro	261
	CAITCCAGTGCATGCAGGGGCTATTGCACCAGGGCCCAATGAGAGAACCCAGGGGAAGTGCATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT	
1108	IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp	301
	ATCCCAAGTASAGAGAAATCTATAAAGATGGATAATCTCTGGGATTAAATATAAATAGTAAGAAATGTATAGCCCTACCAGCATCTGGACATAGACACAGGACCAAGGAAACCCCTTTACAGAT	
1228	TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsnTrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys	341
	TATGTAGACCGGTCTATAAANCTCTAAGAGCCGCAACAGCTTCACAGGATGTAAANATTCGATGACAGAACCTTGTGGTCCAAATGCAAAACCCAGATTGTAAAGACTATTTTAAAT	
1348	AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgValLeuStop	ptac 5
	GCATTGGGACCCAGCAGCTACACTAGAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCCGGCAATAAGCAAGAGTTTTCGATAG	

Figure 10

ARV GAG p16 - synthetic Parts A and B

5' ^{arv 234}
 MetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLys
 TATTATGCAAAGAGGTAACCTTCAGGAATCAAAGAAAGACCGTTAAGTGTTCAACTGTGGTAAG
 ATATACGTTTCTCCATTGAAGTCCTTAGTTTCTTTCTGGCAATTCACAAAGTTGACACCATTCTC
 3' ^{arv 235}
 10 mn11, 23 hinf1, 5'
 63 GluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysAlaCysTrpArgCysGly
 GAAGGTCACATCGCTAAGAACTGTAGAGCTCCAAGAAAGAGGCTTGTGGAGATGTGGT
 CTTCCAGTGTAGCGATTCTTGACATCTCGAGGTTCTTTCTTCCGAACAACCTCTACACCA
 76 dde1, 88 ban2 hgiA hgiJ11 sac1 sduI, 89 alu1,
 123 ArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle
 AGAGAAGGTCACCAAATGAAGGACTGTACCGAAAGACAAGCTAACTTCTTGGGTAAAGATC
 TCTCTTCCAGTGGTTTACTTCTGACATGGCTTCTGTTCGATTGAAGAACCCATTCTAG
 129 bstE2, 131 hph, 148 rsal, 161 alu1, 178 bgl11 xho2, 179
 sau3a,
 183 TrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAla
 TGGCCATCTTACAAGGGTAGACCAGGTAACCTTCTTGCAATCCAGACCAGAACCAACCGCT
 ACCGGTAGAATGTTCCCATCTGGTCCATTGAAGAACGTTAGGTCTGGTCTTGGTTGGCGA
 183 bal1 cfr1 hae1, 184 hae111, 199 acc1, 204 apy1 ecor11 sc
 rF1,
 243 ProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLysGlnGlu
 CCACCTGAAGAAAGTTTTCAGGTTTCGGTGAAGAAAAGACCACCCCATCTCAAAGCAAGAA
 GGTGGACTTCTTTCAAAGTCCAAGCCACTTCTTTTCTGGTGGGTAGAGTTTTCGTTCTT
 249 mbo11, 267 hph, 270 mbo11,
 303 ProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsnAspPro
 CCAATCGACAAGGAATTGTACCCATTGACCTCTTTGAGATCCTTGTTCGGTAACGATCCC
 GGTTAGCTGTTCTTAACATGGGTAACCTGGAGAACTCTAGGAACAAGCCATTGCTAGGG
 307 taq1, 320 rsal, 331 mn11, 339 xho2, 340 sau3a, 357 sau3a
 , 361 mn11, 362 avai1 xho1,
 363 SerSerGlnOP AM
 TCGAGCCAATGATAG
 AGCTCGGTTACTATCAGCT
 363 taq1, 377 acc1 hind11 sal1

Fig. 11

Nucleotide
positions
relative to
FIGURE 5.

MetIleAspLysAlaGlnGluGluHisGluLysTyrHisSerAsnTrp

1 AGGXAAACAG:::ATGAT:GA:AAGGCACAAGAAGAACATGAGAAATATCACAGTAATTGG
TCCXTTGTCT:::TACTA:CT:TTCCGTGTTCTTCTTGTACTCTTTATAGTGTCTTAACC

32 mbo11, 38 nla111,

3820 62 ArgAlaMetAlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSer
AGAGCCATGGCTAGTGATTTTAACTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGC
TCTCGGTACCGATCACTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTCTG

66 nco1, 67 nla111, 118 nspBII pvu11, 119 alu1,

3880 122 CysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGly
TGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGA
ACACTATTTACAGTTCGATTTTCTCTTCGGTACGTACCTGTTTCTCTGACATCAGGTCTCT

135 alu1, 151 nla111, 152 nsi1 ava3, 155 nla111, 164 acc1, 1
76 apy1 bstXI ecor11 scrF1,

3940 182 IleTrpGlnLeuAspCysThrHisLeuGluGlyLysIleIleLeuValAlaValHisVal
ATATGGCAACTAGATTGTACACATCTAGAAGGAAAAATTATCTGGTAGCAGTTTCATGTA
TATACCGTTGATCTAACATGTGTATGATCTTCTTTTAAATAGGACCATCGTCAAGTACAT

198 rsaI, 205 xba1, 223 apy1 ecor11 scrF1, 236 nla111,

4000 242 AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyr
GCCAGTGGATATATAGAAGCAGAAAGTTATTCCAGCAGAGACAGGGCAGGAAACAGCATAT
CGGTACCTATATATCTTCGTCTTCAATAAGGTCGTCTCTGTCCCCTCTTTGTCGTATA

263 xmn1,

4060 302 PheLeuLeuLysLeuAlaGlyArgTrpProValLysThrIleHisThrAspAsnGlySer
TTTCTCTTAAATAGCAGGAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGC
AAAGAGAATTTTAAATCGTCTTCTACCGGTCAATTTTGTATGTATGTCTGTTACCGTCTG

321 mbo11, 326 bal1 cfr1 hae1, 327 hae111, 357 bbv fnu4h1,

4120 362 AsnPheThrSerThrThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPhe
AATTTCAACAGTACTACGGTTAAAGGCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTT
TTAAAGTGGTTCATGATGCCAATTCCGGCGGACAACCACCCGTCCTAGTTTCGTCTTAAA

366 hph, 371 sca1, 372 rsaI, 385 hae111, 386 fnu4h1 nsb11, 4
05 binI, 406 dpn1 sau3a,

4180 422 GlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMetAsnAsnGluLeuLys
GGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAAG
CCGTAAGGGATGTTAGGGGTTTCAGTTCTCATCATCTTAGATACTTATTACTTAATTTCT

423 bsm1, 458 hinf1,

4240 482 LysIleIleGlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAla
AAAATTATAGGACAGGTAAGAGATCAGGCTGAACACCTTAAAGACAGCAGTACAAATGGCA
TTTTAATATCCTGTCCATTCTCTAGTCCGACTTGTGGAATTCTGTCTCATGTTTACCGT

503 dpn1 sau3a, 518 afl111, 530 rsaI,

4300 542 ValPheIleHisAsnPheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArg
GTATTCATCCACAATTTTAAAGAAAAGGGGGGATTGGGGGATACAGTGCAGGGGAAAGA
CATAAGTAGGTGTTAAATTTTCTTTTCCCCCTAACCCTATGTCACGTCCCCTTTCT

547 fok1, 557 aha111,

4360 602 IleValAspIleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLys
ATAGTAGACATAATAGCAACAGACATACAACTAAAGAACTACAAAAGCAAATTACAAAA
TATCATCTGTATTATCGTTGTCTGTATGTTTGATTCTTGATGTTTTCTGTTTAAATGTTTT

605 acc1,

662 IleGlnAsnPheArgValTyrTyrArgAspAsnLysAspProLeuTrpLysGlyProAla
ATTCAAAATTTTTCGGGTTTATTACAGGGACAACAAAGATCCCCTTTGGAAAGGACCAGCA

4480

722

LysLeuLeuIrpLysGivGluGlyAlaValValIleGlnAspAsnSerAspIleLysVal
 AAGCTTCTCTGGAAAGGTGAAGGGGCGAGTAGTAATACAAGATAATAGTGACATAAAAGTA
 TTCGAAGAGACCTTTCCACTTCCCGTCATCATTATGTTCTATTATCACTGTATTTTCAT

722 hind111, 723 alu1, 737 hph,

4540

782

ValProArgArgLysAlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAsp
 GTGCCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGAT
 CACGGTTCTTCTTTTCGTTTTTAGTAATCCCTAATACCTTTTGTCTACCGTCCACTACTA

789 mbo11, 833 hph,

4600

842

CysValAlaSerArgGlnAspGluAspAM
 TGTGTGGCAAGTAGACAGGATGAGGATTAGTCGACGGAATTCTTTAGTAAACACC
 ACACACCGTTCATCTGTCCTACTCCTAATCAGCTGCCTTAAGAAATCATTTTGTGG

852 acc1, 859 fok1, 863 mn11, 871 acc1 hind11 sal1, 872 taq1
 , 878 ecor1,

FIGURE 12

2 of 2

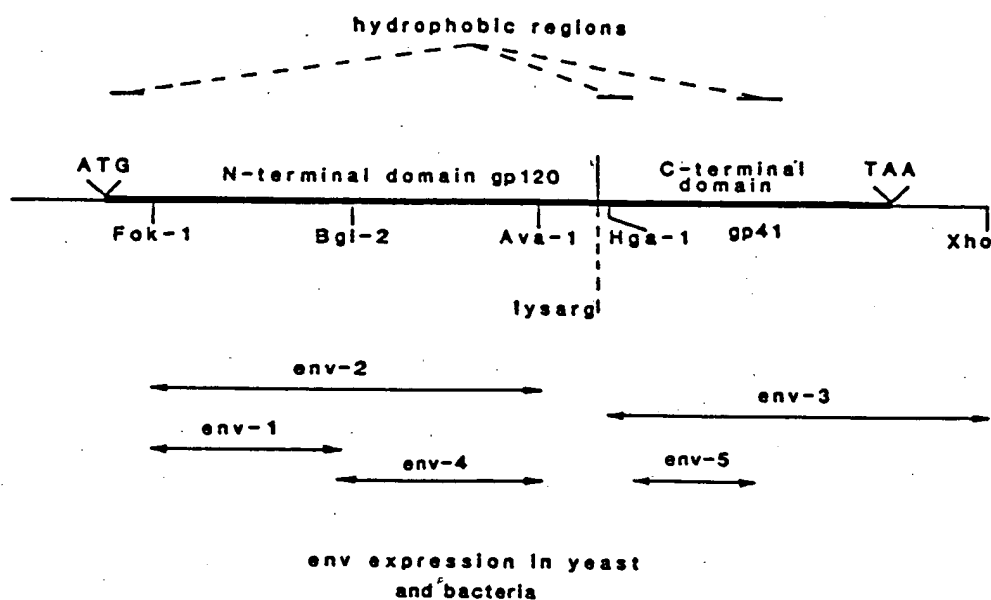


FIGURE 13

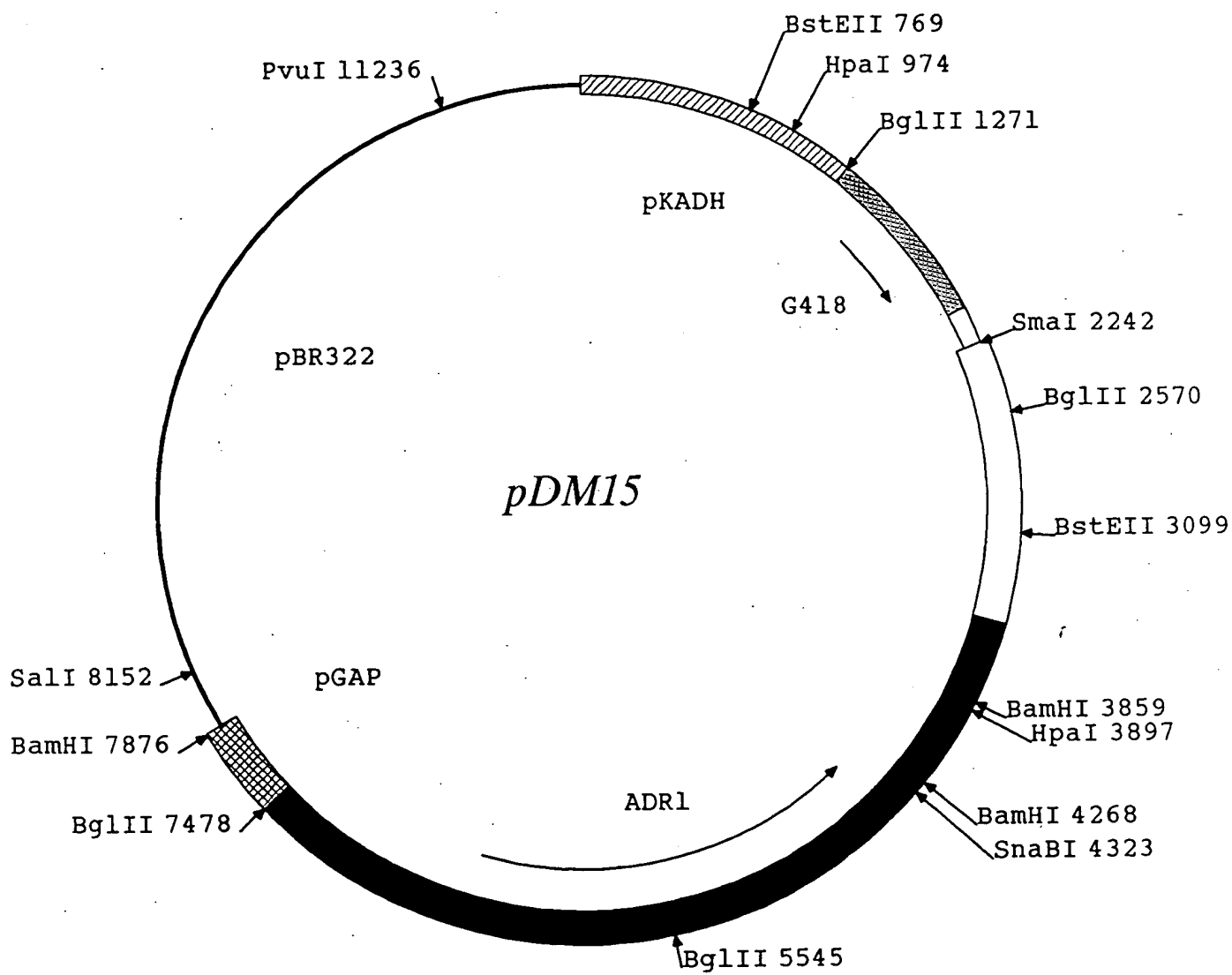


FIGURE 14

SCD
MetAlaThrLysAlaValCysValLeuLysGlyAspGlyProValGlnGlyIleIleAsn
1 CATGGCGACGAAGGCCGTGTGCGTGCCTGAAGGGCGACGGCCAGTGCAGGGCATCATCAAT
CGCTGCTTCCGGCACACGCACGACTTCCCCGCTGCCGGGTACAGTCCCCGTAGTAGTTA

62 PheGluGlnLysGluSerAsnGlyProValLysValTrpGlySerIleLysGlyLeuThr
TTCGAGCAGAAGGAAAGTAATGGACCACTGAAGGTGTGGGGAAGCATTAAAGGACTGACT
AAGCTCGTCTTCTTTCATTACCTGGTCACTTCCACACCCCTTCGTAATTTCTGACTGA

122 GluGlyLeuHisGlyPheHisValHisGluPheGlyAspAsnThrAlaGlyCysThrSer
GAAGGCCCTGCATGGATTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGT
CTTCCGGACGTACCTAAGGTACAAGTACTCAAACCTCTATTATGTCTGCCACATGGTCA

182 AlaGlyProHisPheAsnProLeuSerArgLysHisGlyGlyProLysAspGluGluArg
GCAGGTCTCTCACTTTAATCTCTATCCAGAAAACACGGTGGGCCAAAGGATGAAGAGAGG
CGTCCAGGAGTGAATTTAGGAGATAGGTCTTTTGTGCCACCCGGTTTCTACTTCTCTCC

242 HisValGlyAspLeuGlyAsnValThrAlaAspLysAspGlyValAlaAspValSerIle
CATGTTGGAGACTTGGGCAATGTGACTGCTGACAAAGATGGTGTGGCCGATGTGTCTATT
GTACAACCTCTGAACCCGTTACACTGACGACTGTTTCTACCACACCCGGCTACACAGATAA

302 GluAspSerValIleSerLeuSerGlyAspHisCysIleIleGlyArgThrLeuValVal
GAAGATTCTGTGATCTCACTCTCAGGAGACCATTTGCATCATTTGGCCGCACACTGGTGGTC
CTTCTAAGACACTAGAGTGAGAGTCTCTGGTAACGTAGTAACCGGCGTGTGACCACCAG

362 HisGluLysAlaAspAspLeuGlyLysGlyGlyAsnGluGluSerThrLysThrGlyAsn
CATGAAAAGCAGATGACTTGGGCAAGGTGGAAATGAAGAAAGTACAAAGACAGGAAAC
GTACTTTTTCTGCTCTACTGAACCCGTTTCCACCTTTACTTCTTTCTATGTTTCTGTCTTTG

ENV 5B
422 AlaGlySerArgLeuAlaCysGlyValIleGlyIleAlaMetAlaIleGluAlaGlnGln
GCTGGAAGTCGTTTGGCTTGTGGTGTAAATTGGGATCGCCATGGCTATCGAAGCTCAACAA
CGACCTTCAGCAAACCGAACACCATTAACCTAGCGGTACCGATAGCTTCGAGTTGTT

482 HisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaVal
CACTTGCTGCAGTTGACCGTTTGGGGTATCAAGCAGTTGCAGGCTAGAGTTTTGGCTGTT
GTGAACGACGTCAACTGGCAAACCCCATAGTTTCGTCAACGTCCGATCTCAAACCGACAA

542 GluArgTyrLeuArgAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIle
GAAAGATACTTGAGAGATCAACAATTGTTGGGTATCTGGGGTTGTTCTGGTAAGTTGATT
CTTTCTATGAACCTCTCTAGTTGTTAACAACCCATAGACCCCAACAAGACCATTCAACTAA

602 CysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluAspIleTrp
TGTACCACCGCTGTTCCCTGGAACGCTTCTTGGTCTAACAAGTCTTTGGAAGACATCTGG
ACATGGTGGCGACAAGGGACCTTGCGAAGAACCAGATTGTTTCAGAAACCTTCTGTAGACC

662 AspAsnMetThrTrpMetGlnTrpGluArgGluIleAspAsnTyrThrAsnThrIleTyr
GACAACATGACCTGGATGCAATGGGAAAGAGAAATCGACAACCTACACCAACACCATCTAC
CTGTTGTACTGGACCTACGTTACCTTTCTCTTTAGCTGTGTGATGTGGTGTGGTAGATG

722 ThrLeuLeuGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu
ACCTTGTTGGAGGAATCTCAAACCAACAAGAAAAGAACGAACAAGAATTGTTGGAATTG
TGGAACAACCTCCTTAGAGTTTTGGTTGTTCTTTTCTTGCTGTCTTACAACCTTAAC

782 AspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrAsnTrpAM
GACAAGTGGGCAAGCTTGTGGAAGTGGTCTCTATCACCACCTGGTAG
CTGTTCAACCGTTTGAACACCTTGACCAAGAGATAGTGGTTGACCATCAGCT

Translated Mol. Weight = 30414.22

FIGURE 15

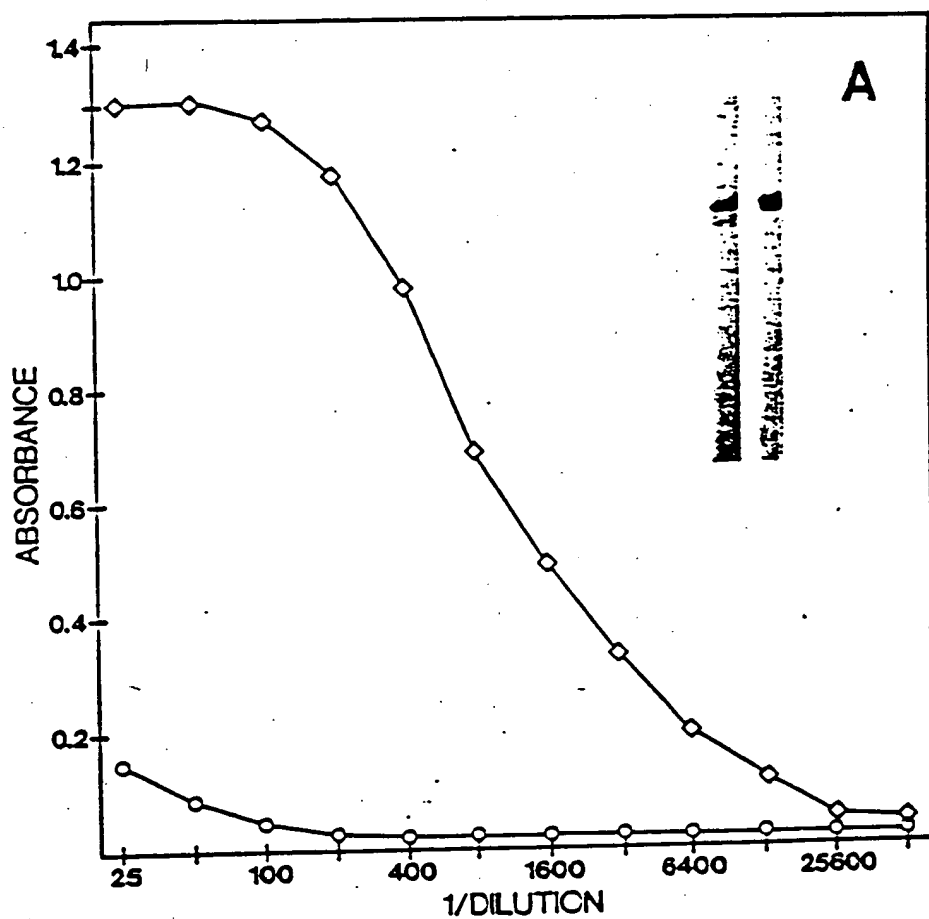


FIGURE 16

2 of 2

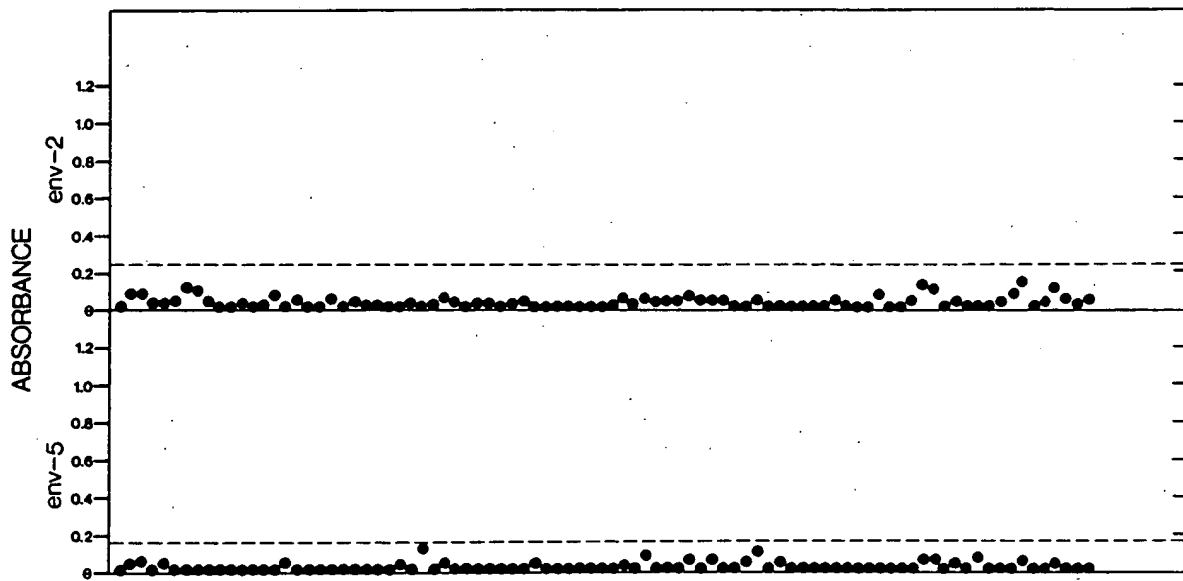


FIGURE 17

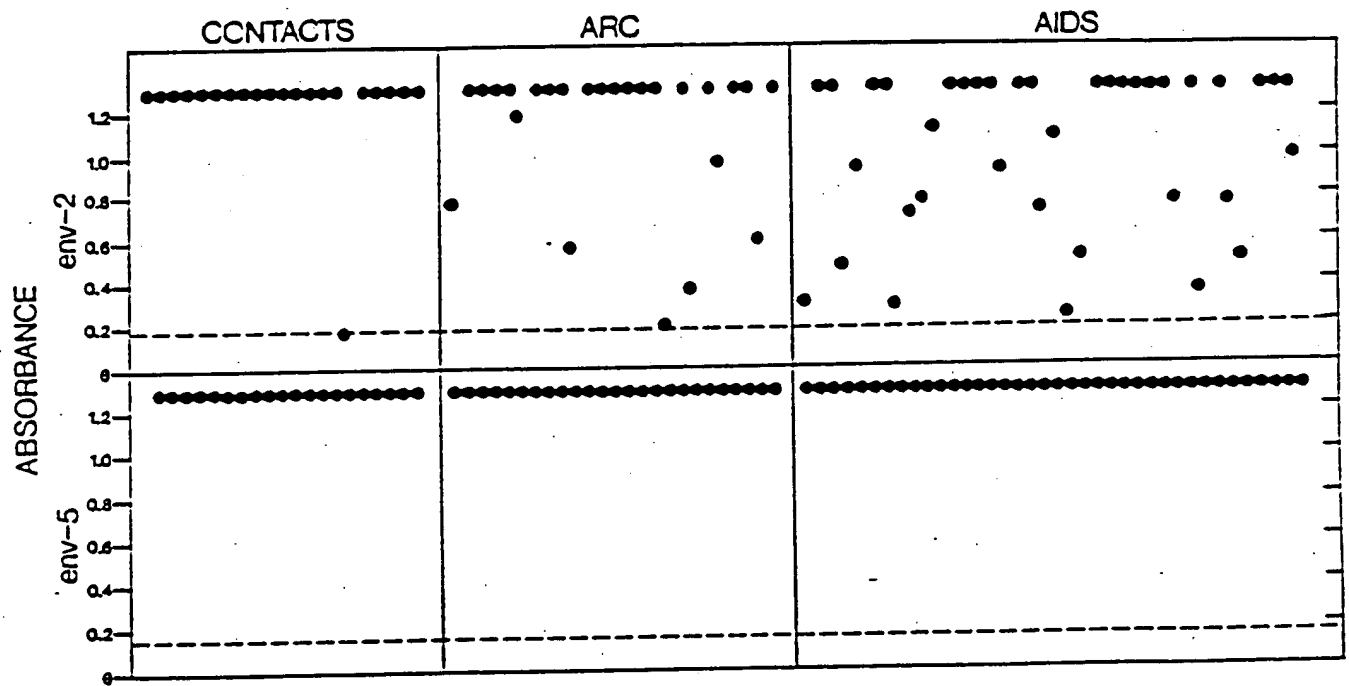


FIGURE 18

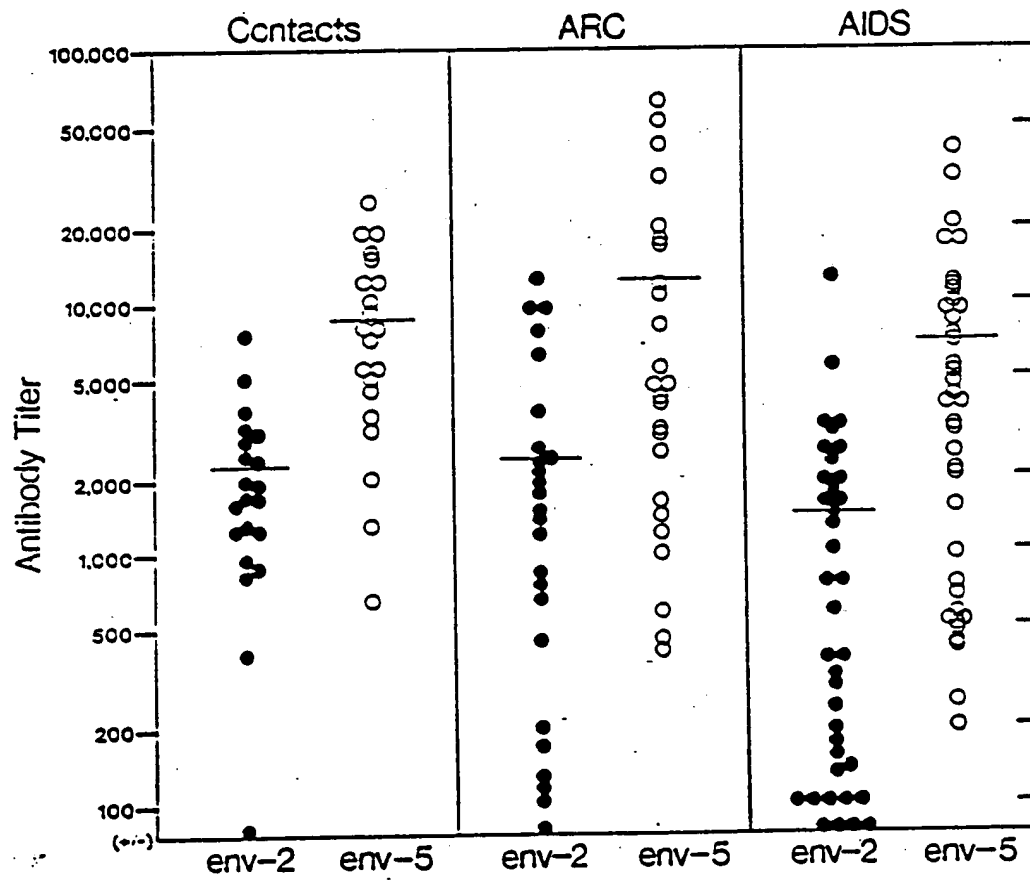


FIGURE 19

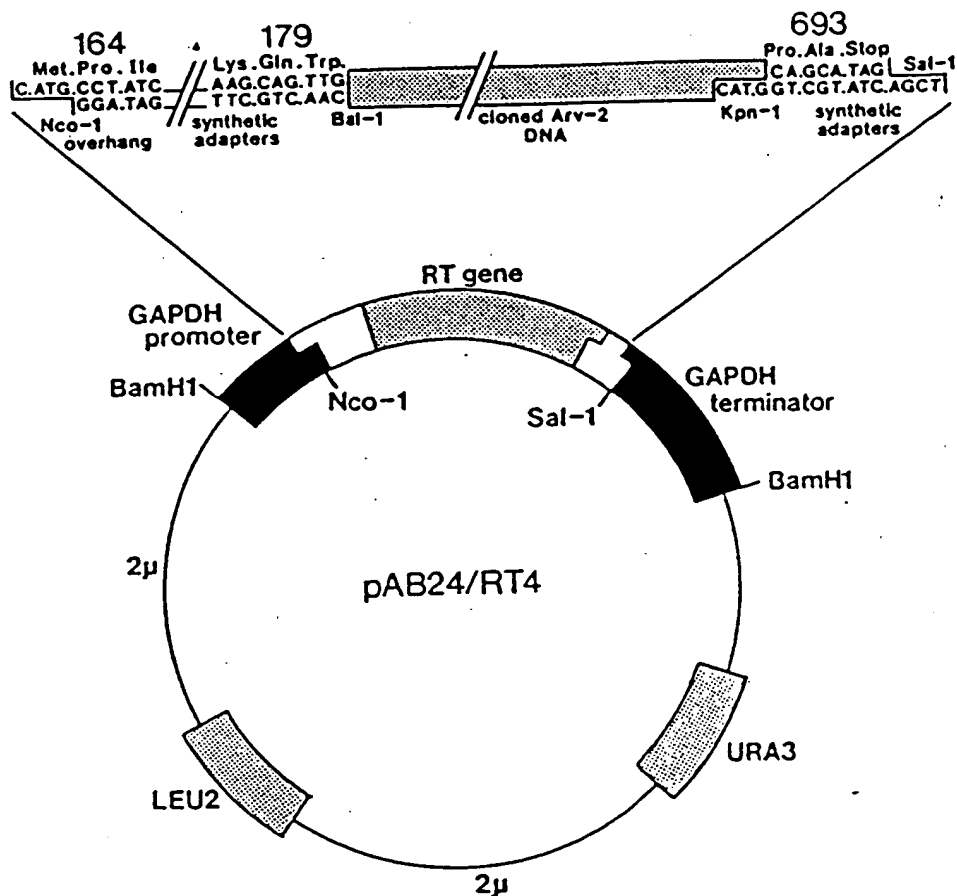


FIGURE 20

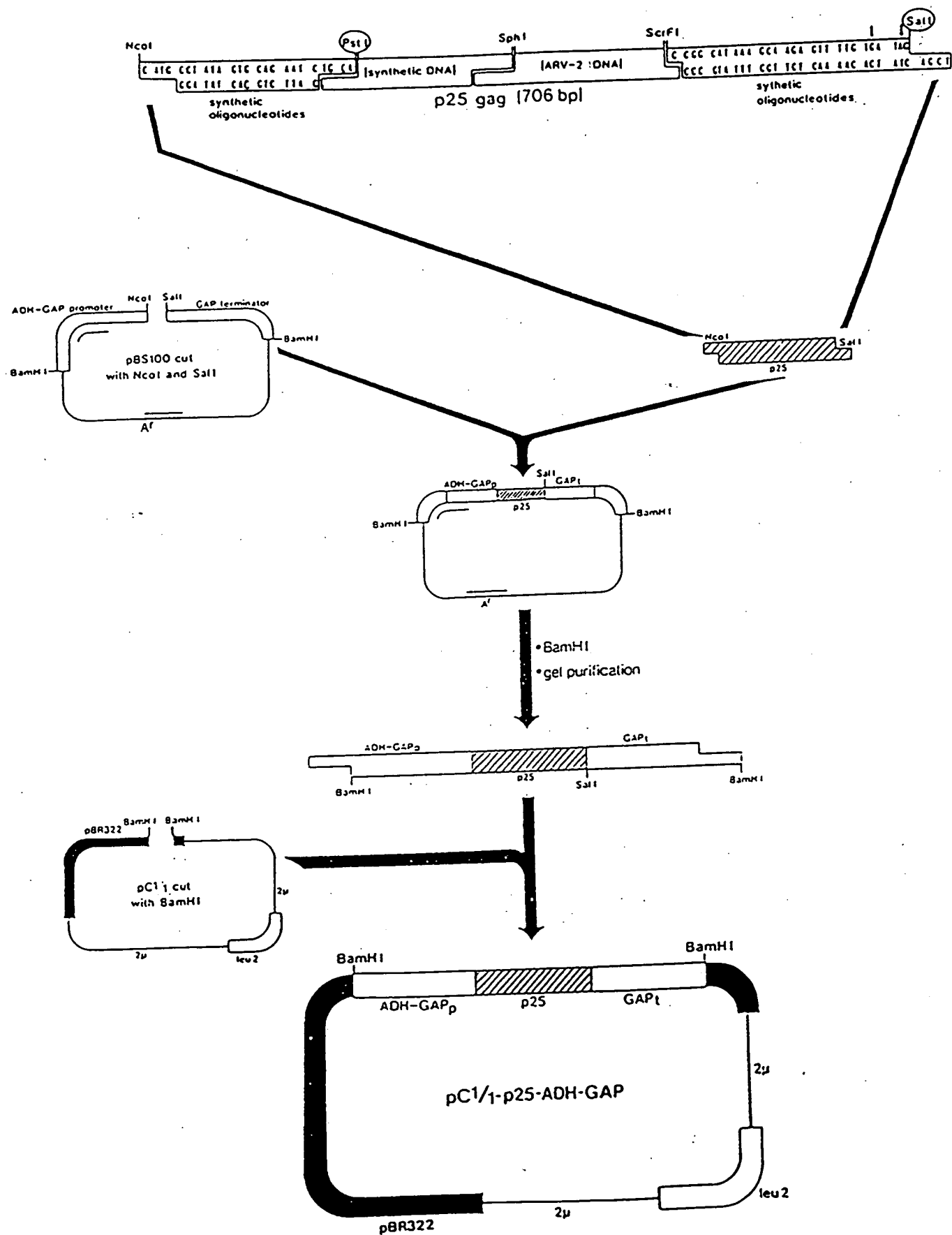


FIGURE 21

```

      1      10
Met Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln
C ATG CCT ATA GTG CAG AAT CTG CAG GGG CAA ATG GTA CAT CAG

      20
Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
GCC ATA TCA CCT AGA ACT TTA AAT GCT TGG GTA AAA GTA GTA GAA

      30      40
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu
GAA AAG CCT TTC AGC CCA GAA GTA ATA CCC ATG TTT TCA GCA TTA

      50
Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr
TCA GAA GGA GCC ACC CCT CAA GAT TTA AAC ACC ATG CTA AAC ACA

      60      70
Val Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile
GTG GGG GGA CAT CAA GCA GCC ATG CAA ATG TTA AAA GAG ACT ATC

      80
Asn Glu Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala
AAT GAG GAG GCT GCC GAA TGG GAT AGA GTG CAT CCA GTG CAT GCA

      90      100
Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp
GGG CCT ATT GCA CCA GGC CAA ATG AGA GAA CCA AGG GGA AGT GAC

      110
Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met
ATA GCA GGA ACT ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG

      120      130
Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp
ACA AAT AAT CCA CCT ATC CCA GTA GGA GAA ATC TAT AAA AGA TGG

      140
Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr
ATA ATC CTG GGA TTA AAT AAA ATA GTA AGA ATG TAT AGC CCT ACC

      150      160
Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp
AGC ATT CTG GAG ATA AGA CAA GGA CCA AAG GAA CCC TTT AGA GAT

      170
Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln Ala Ser
TAT GTA GAC CGG TTC TAT AAA ACT CTA AGA GCC GAA CAA GCT TCA

      180      190
Gln Asp Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val Gln Asn
CAG GAT GTA AAA AAT TGG ATG ACA GAA ACC TTG TTG GTC CAA AAT

      200
Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Ala
GCA AAC CCA GAT TGT AAG ACT ATT TTA AAA GCA TTG GGA CCA GCA

      210      220
Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG GCA GTG GGG GGA

      230      232
Pro Gly His Lys Ala Arg Val Leu OP
CCC GGG CAT AAA GCA AGA GTT TTG TGA TAG

```

Translated Mol. Weight = 25700.75

FIGURE 22

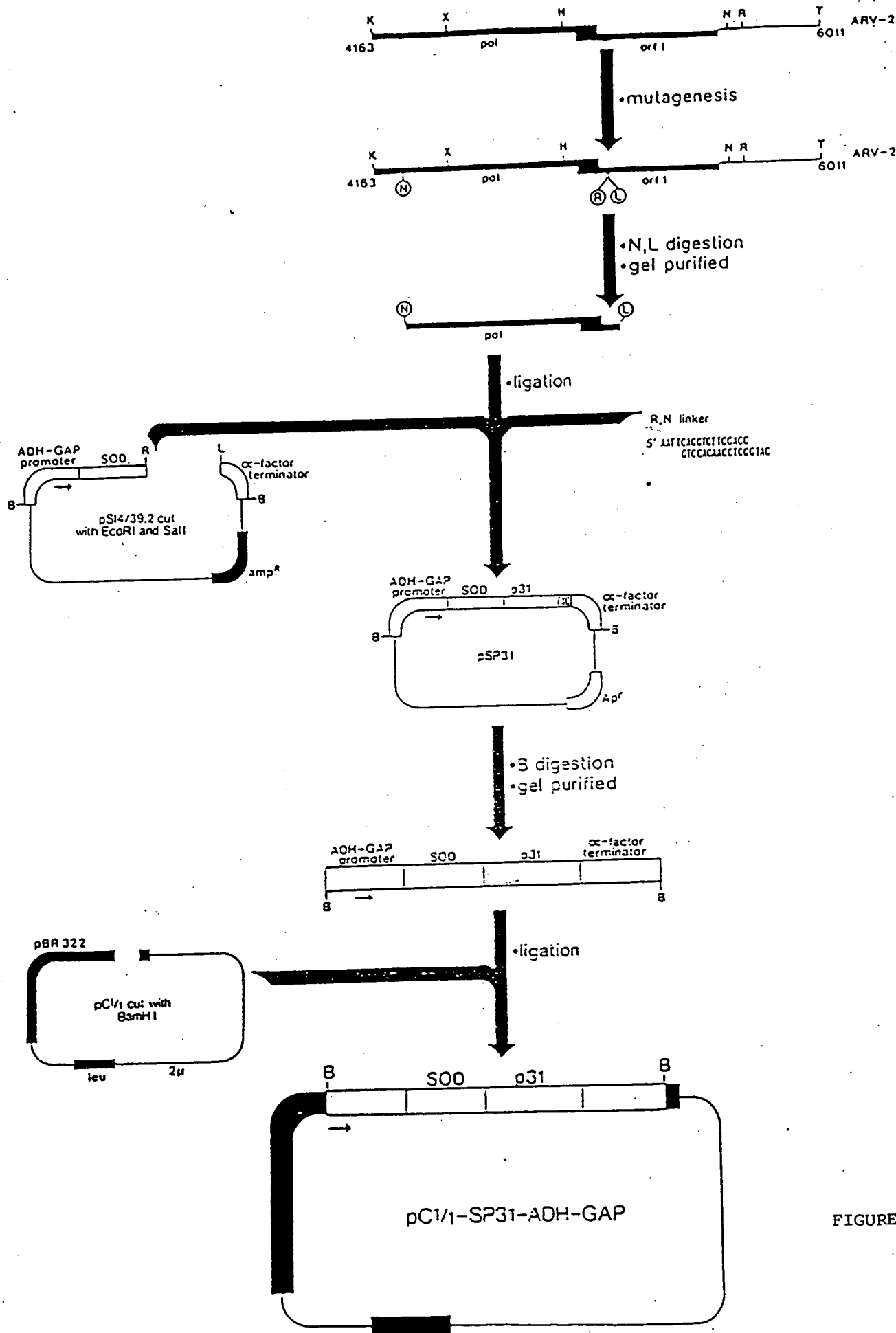


FIGURE 23

1383 ValCysValLeuLysGlyAspGlyProValGlnGlyIleIleAsnPheGluGlnLysGlu
GTTTGTGTTTTGAAGGGTGACGGCCAGTTCAAGGTATTATTAACTTCGAGCAGAAGGAA
CAACACAAAACCTCCCACTGCCGGGTCAAGTTCCATAATAATTGAAGCTCGTCTTCCTT

1443 SerAsnGlyProValLysValTrpGlySerIleLysGlyLeuThrGluGlyLeuHisGly
AGTAATGGACCAGTGAAGGTGTGGGGAAGCATTAAAGGACTGACTGAAGGECCTGCATGGA
TCATTACCTGGTCACTTCCACACCCCTTCGTAATTTCTGACTGACTTCCGGACGTACCT

1503 PheHisValHisGluPheGlyAspAsnThrAlaGlyCysThrSerAlaGlyProHisPhe
TTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGTGCAGGTCCCTCACITTT
AAGGTACAAGTACTCAACCTCTATTATGTCTGCCGACATGGTCACGTCCAGGAGTGAAA

1563 AsnProLeuSerArgLysHisGlyGlyProLysAspGluGluArgHisValGlyAspLeu
AATCTCTATCCAGAAAACACGGTGGGCCAAAGGATGAAGAGAGGCGATGTTGGAGACTTG
TTAGGAGATAGGTCCTTTGTGCCACCCGGTTCTACTTCTCTCCGTACAACCTCTGAAC

1623 GlyAsnValThrAlaAspLysAspGlyValAlaAspValSerIleGluAspSerValIle
GGCAATGTGACTGCTGACAAAGATGGTGTGGCCGATGTGTCTATTGAAGATTCTGTGATC
CCGTTACACTGACGACTGTTTCTACACACCCGGTACACAGATAACTTCTAAGACACTAG

1683 SerLeuSerGlyAspHisCysIleIleGlyArgThrLeuValValHisGluLysAlaAsp
TCACTCTCAGGAGACCATTGTCATCATTGGCCGCACACTGGTGGTCCATGAAAAGCAGAT
AGTGAGAGTCTCTGGTAACGTAGTAACCGGCGTGTGACCACCAGGTACTTTTTCGTCTA

1743 AspLeuGlyLysGlyGlyAsnGluGluSerThrLysThrGlyAsnAlaGlySerArgLeu
GACTTGGGCAAGGTGGAAATGAAGAAAGTACAAAGACAGGAAACGCTGGAAGTCGTTTG
CTGAACCCGTTTCCACCTTTACTTCTTTCATGTTTCTGTCTTTGCGACCTTCAGCAAAC

1803 AlaCysGlyValIleGlyIleAlaGlnAsnSerGlyValGlyAlaMetAlaMetAlaSer
GCTTGTGGTGTAAATTTGGGATCGCCAGAAATCAGGTGTTGGAGCCATGGCCATGGCTAGT
CGAACACCACATTAACCTAGCGGGTCTTAAGTCCACAACCTCGGTACCGGTACCGATCA

1863 AspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLysCysGln
GATTTTAACTTCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAATGTGAG
CTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTGACACTATTTACAGTC

1923 LeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAsp
CTAAAAGGAGAAGCCATGCATGGACAGTAGACTGTAGTCCAGGAATATGGCACTAGAT
GATTTTCTCTCTCGGTACGTACCTGTTTCATCTGACATCAGGTCTTATACGTTGATCTA

1983 CysThrHisLeuGluGlyLysIleIleLeuValAlaValHisValAlaSerGlyTyrIle
TGTACACATCTAGAAGGAAAAATTATCTGGTAGCAGTTCATGTAGCCAGTGGATATATA
ACATGTGTAGATCTTCTTTTAAATAGGACCATCGTCAAGTACATCGGTACCTATATAT

2043 GluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeu
GAAGCAGAAGTTATTTCCAGCAGAGACAGGGCAGGAAACAGCATATTTTCTCTTAAATTA
CTTCGTCTTCAATAAGGTCTCTCTGTCCCGTCTTTGTGCTATAAAAGAGAATTTTAAT

2103 AlaGlyArgTrpProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerThr
GCAGGAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGCAATTTACCAGTACT
CGTCCTTCTACCGGTCAATTTTGTATTGTATGTCTGTTACCGTCTTAAAGTGGTCATGA

2163 ThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsn
ACGGTTAAGGCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAAT
TGCCAATTCCGGCGGACAACCCCGTCCCTAGTTTCTGCTTAAACCGTAAGGGATGTTA

2223 ProGlnSerGlnGlyValValGluSerMetAsnAsnGluLeuLysLysIleIleGlyGln
CCCCAAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAAGAAAATTATAGGACAG
GGGGTTTCAGTTCCTCATCATCTTAGATACTTATTACTTAATTTCTTTTAAATATCCTGTC

2283 ValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIleHisAsn
GTAAAGAGTACGGCTGAACACCTTAAGACAGCAGTACAAATGGCAGTATTTCATCCACAAT
CATTCTCTAGTCCGACTTGTGGAATTCTGTCTCATGTTTACCGTCATAAGTAGGTGTTA

2343 PheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleIle
TTTAAAAGAAAAGGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAGTAGACATAATA
AAATTTTCTTTTCCCCCTAACCCCTATGTCAAGTCCCTTTCTTATCATCTGTATTAT

2403 AlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArg
GCAACAGACATACAACTAAAGAACTACAAAAGCAATTAACAAAATTTTAAAGTTTAAAGGCC
CGTGTCTGTATGTTGATTCTGTATGTTTTCTGTTAATGTTTAAAGTTTAAAGGCC

2463 ValTyrTyrArgAspAsnLysAspProLeuTrpLysGlyProAlaLysLeuLeuTrpLys
GTTTATTACAGGGACAACAAAGATCCCTTTGGAAAGGACCAGCAAGGCTTCTCTGGAAG
CAATAATGTCCCTGTTGTTTCTAGGGGAAACCTTCTCGGTCTGTTTGAAGAGACCTTT

2523 GlyGluGlyAlaValValIleGlnAspAsnSerAspIleLysValValProArgArgLys
GGTGAAGGGGSCAGTAGTAATACAAGATAATAGTGACATAAAGTAGTGCCAAGAAGAAAA
CCACTTCCCCGTATCATATTGTTCTATTATCACGTATTTTTCATCACGGTTCCTCTTTT

2583 AlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArg
GCAAAAATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGA
CGTTTTTAGTAATCCCTAATACCTTTTGTCTACCGTCCACTACTAACACACCGTTTCATCT

2643 GlnAspGluAspAsn
CAGGATGAGGATTAG
GTCCTACTCCTAATC

FIGURE 24

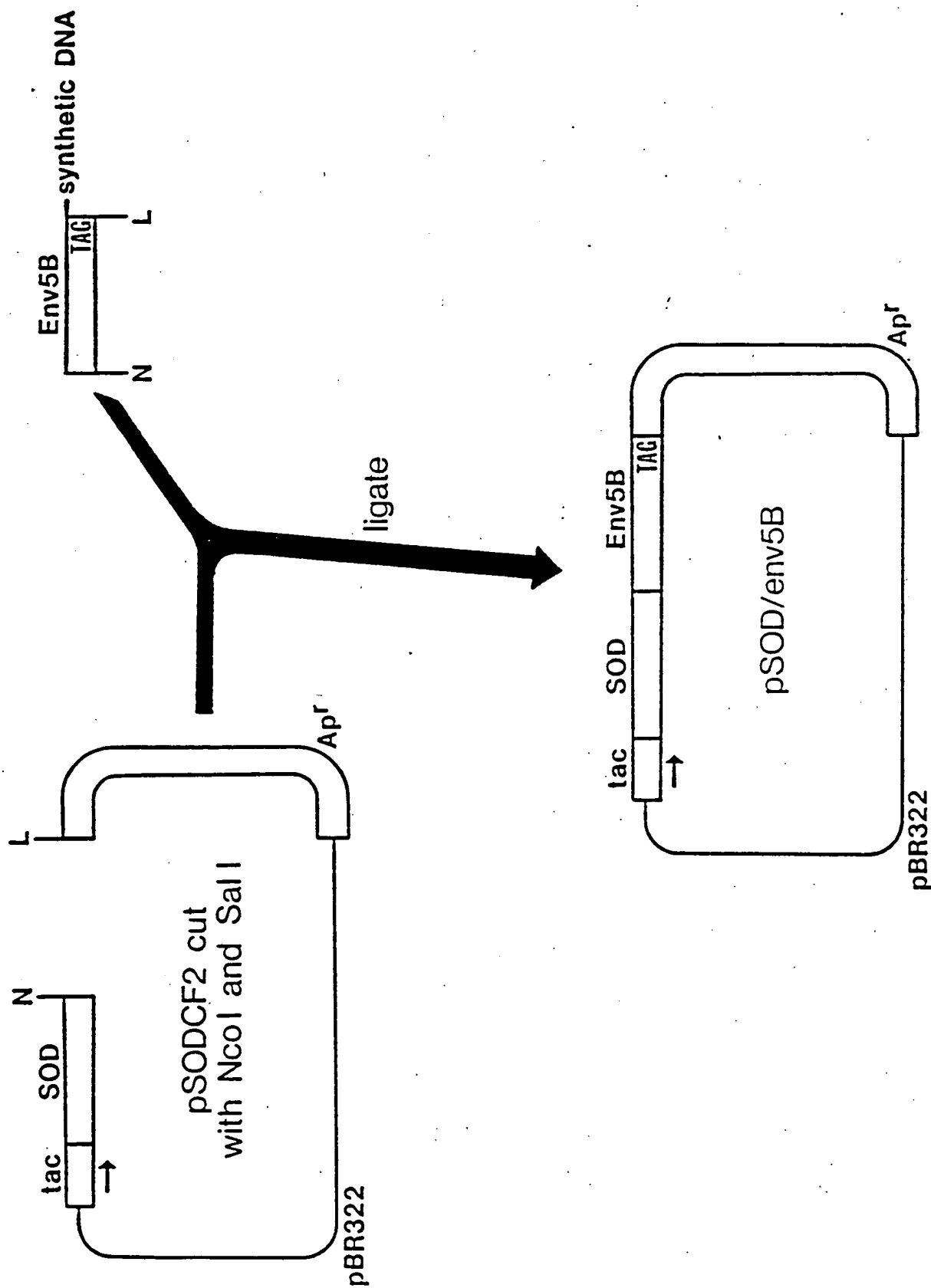


FIGURE 25

Sequence of SOD/env-4

SOD \rightarrow

[illegible]

902	LysGlnIleIleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGly AARCAATTAATAACATATGGCGAGAGTAGAAAGCAATGTATGTCCCTCCCATTTGGA TTTGTATTATTTTGTACACCGCTCTTCATCTCTTTCGTTTACATACGCGGGAGGTAAACCT
962	GlyGlnIleSerCysSerSerAsnIleThrGlyLeuLeuThrArgAspGlyGlyThr GGCAAAATTAGTTGTTCATCAAAATTAATACAGGCTGCTTAATTAACAAGAGATCGTGAC CCGTGTTAATCAACAAGACTAGTTTATTAATGTCCGAGGATTAATTTGTCTCTACCAACCAATGT
1022	AsnValThrAsnAspThrGluValPheArgProGlyGlyGlyAspMetArgAspAsnTrp TAATGAACATAATGACACCGAGCTCTACAGCTCGAGGAGGAGTATTAGGAGCAATGTG TTTACATGTAATTACTTGTCCTCCAGGACTCTGGACCTCTCCTCTACTACTCCCTCTTAAC
1082	ArgSerGluLeuTyrLysValIleLysIleGluProLeuGlyIleAlaProThr AGAAATGCAATTATTAATAATTAAGTATTAATAATTAAGCAATTAGGAATAGACCCACC TCTTCACCTAATATTAATTAATTTCTAATTAATTTTAACTTGTGTAACTCTTATTCGTGTGGTGG
1142	LysAlaLysArgArgValGlnArgGluLysArgOP OP AAGCAAGAGAGAGAGTGGTCGACAGAGAGAAAGATGATGAAGCTTG TTCGGTTCTCTCTACACAGCTCTCTCTTTTCTACTACTACTCTGGAACAGCT

[ysA]a[ysArgArgVa]Va[G]nArgGluLysArgOp Op

1142 AAGGCAAGAGAAGAGTGGTGCAGAGAGAGAAAAAAGATGATGAAGCTTG
TTCCGTTTCTCTCAACCAGCTCTCTCTTTTCTACTACTTGGACAGCT

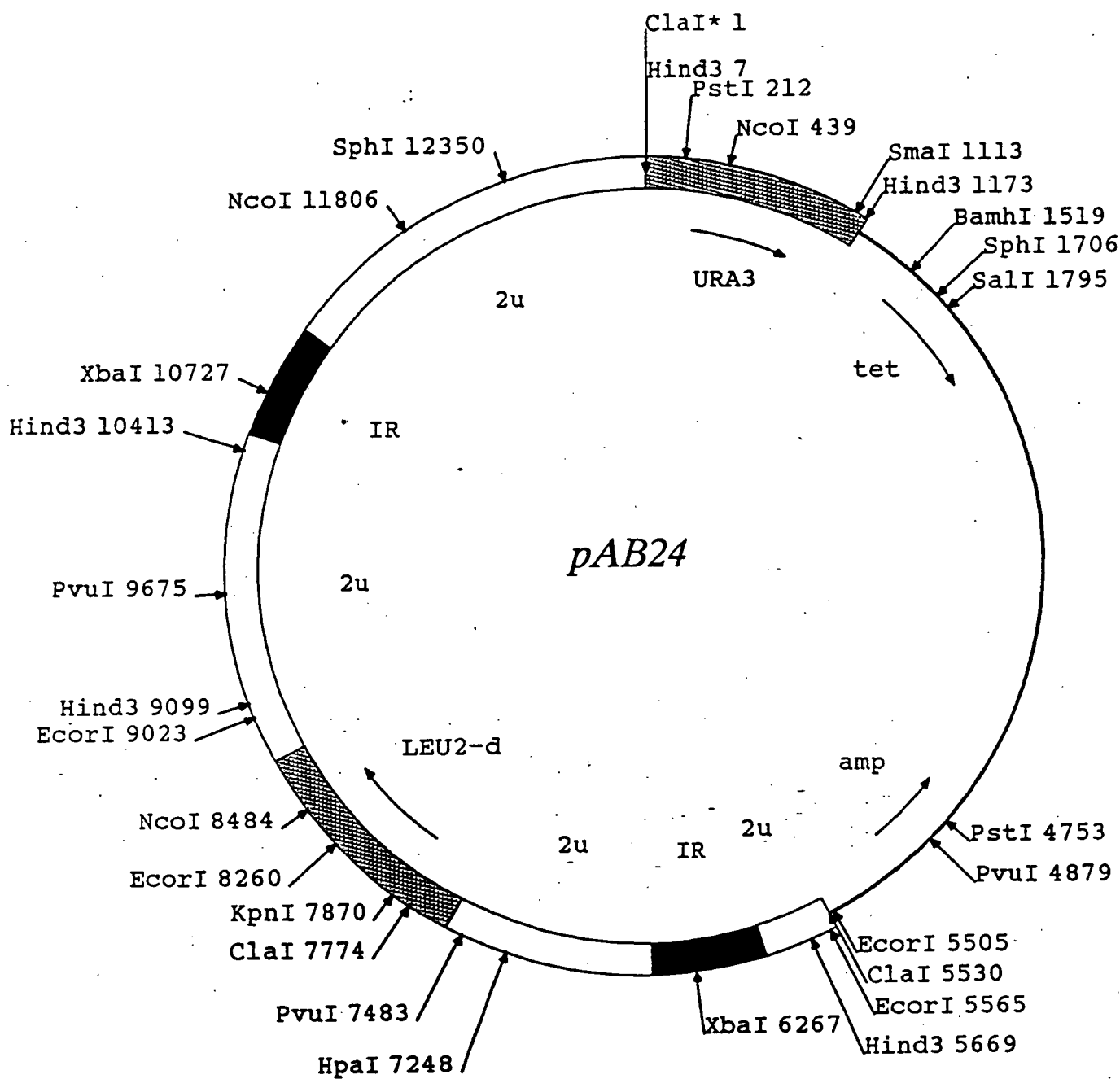


FIGURE 27

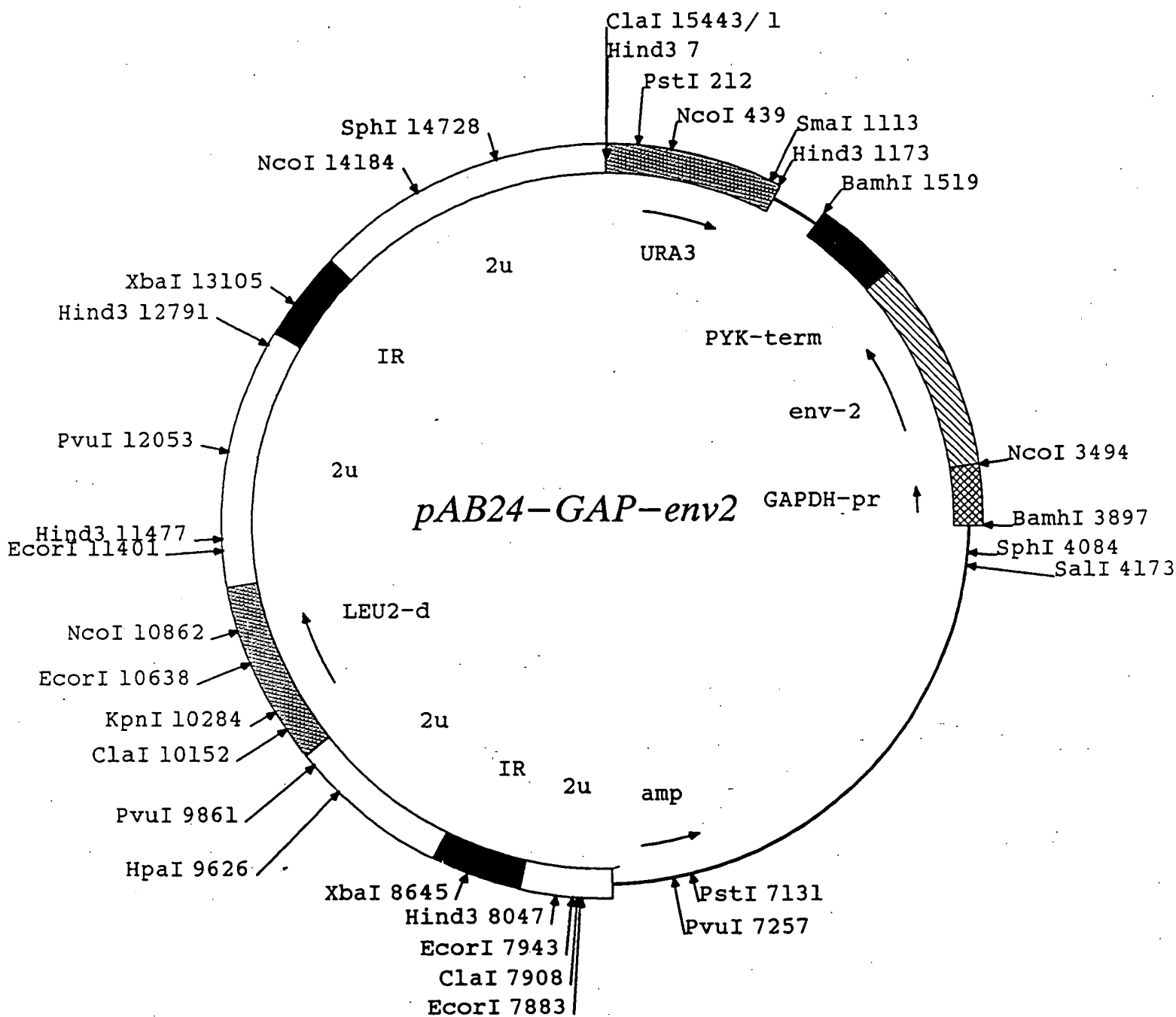


FIGURE 28

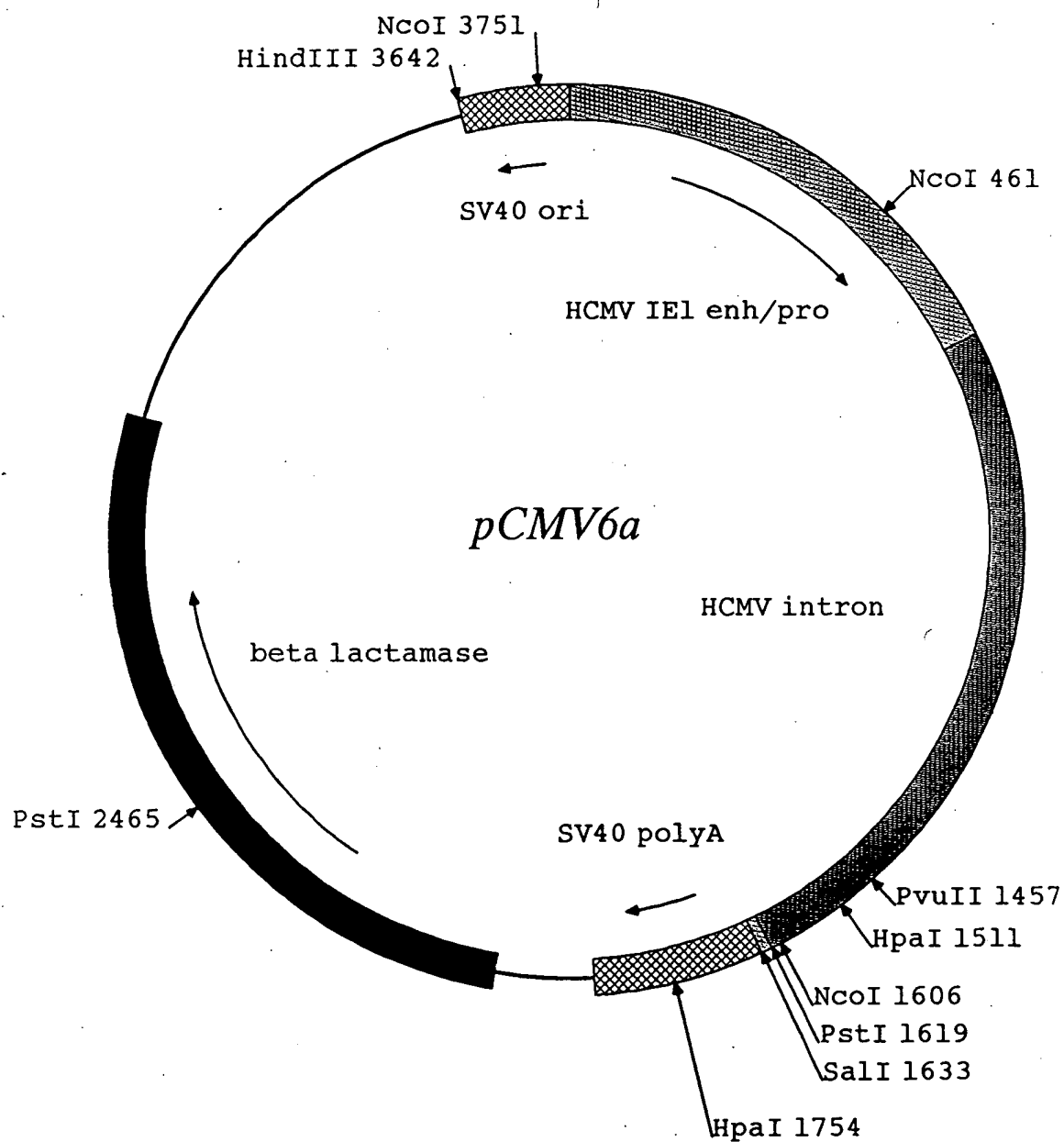


FIGURE 29

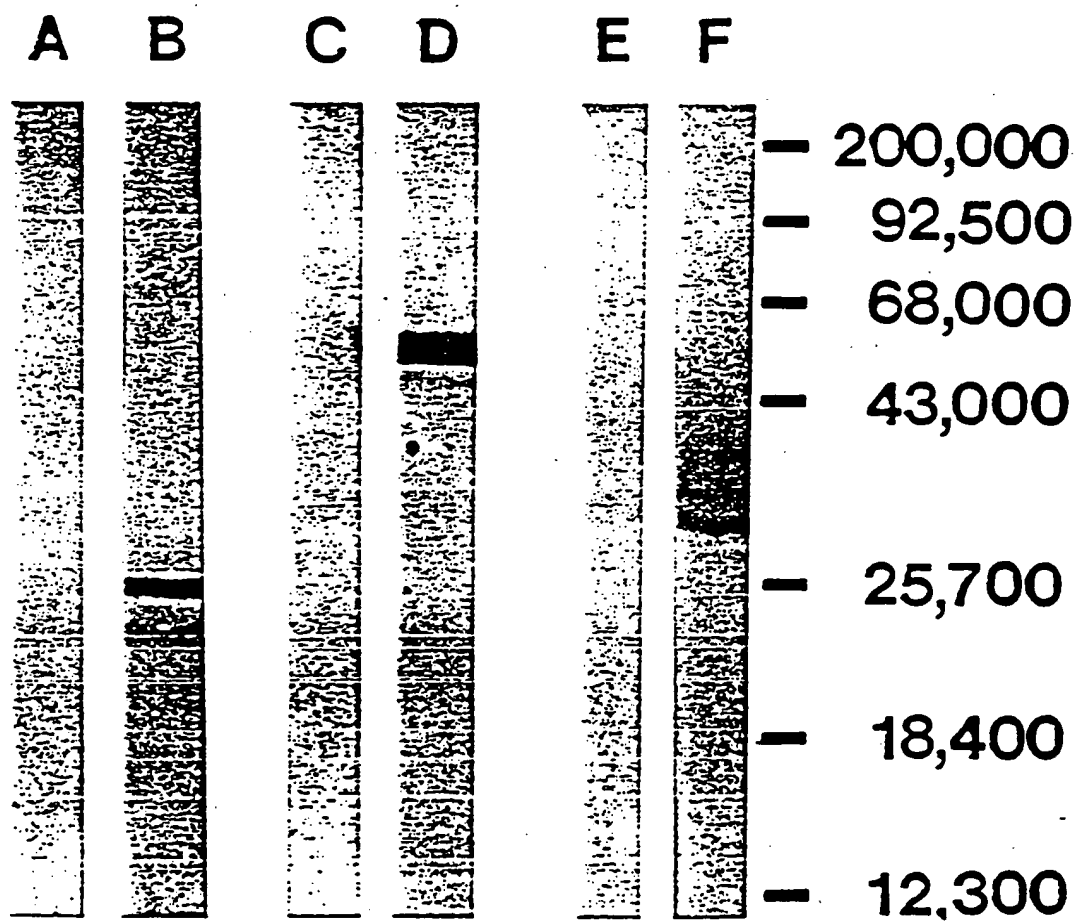


FIGURE 30

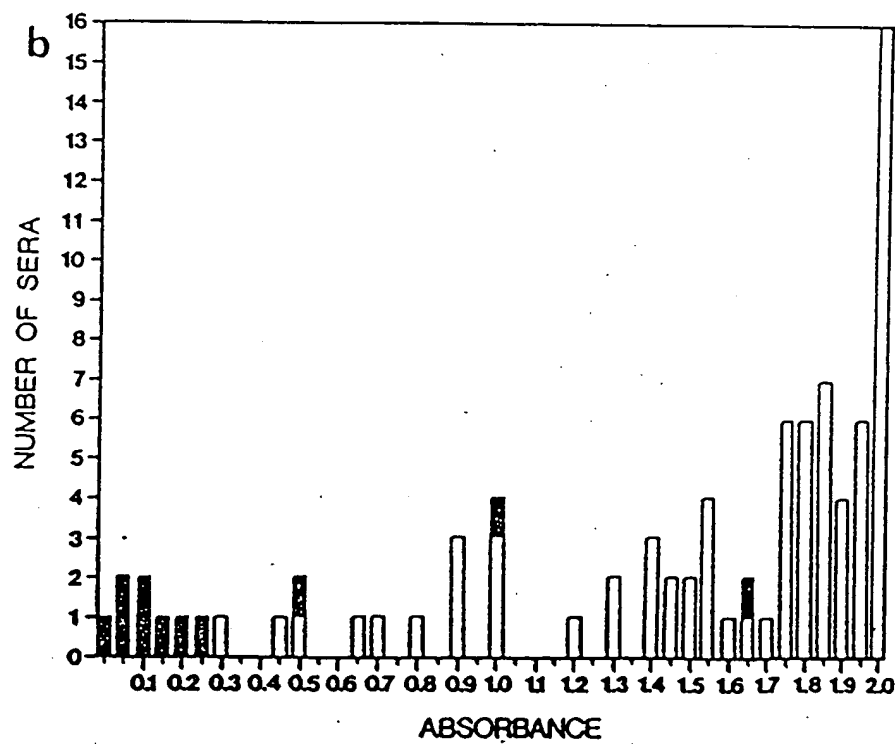
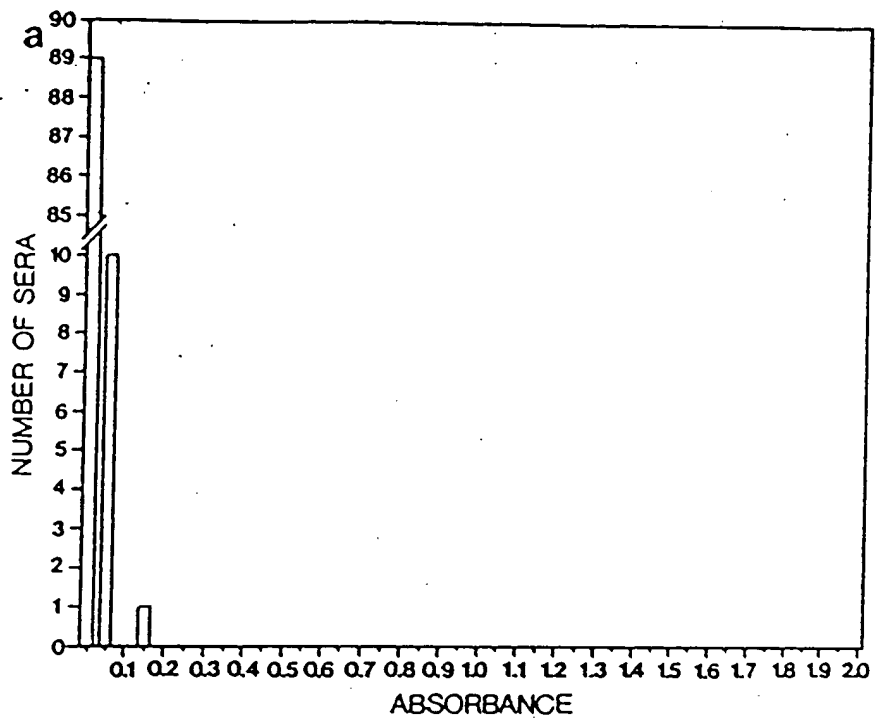


FIGURE 31

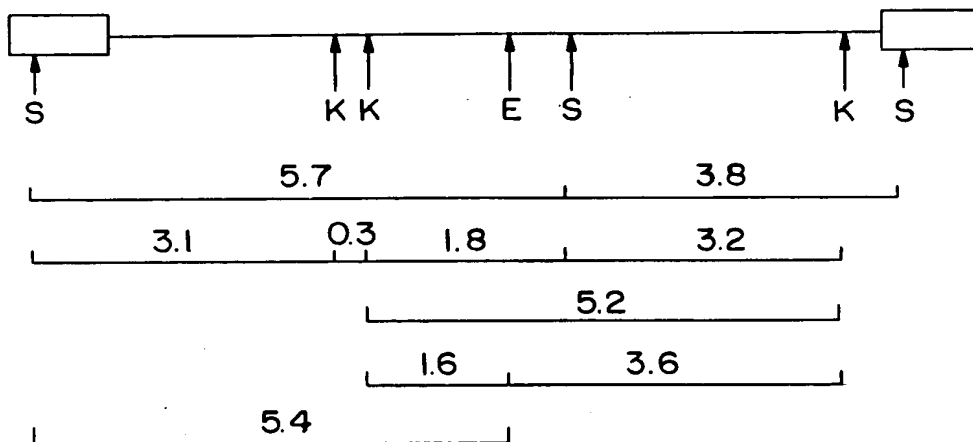


FIG.1

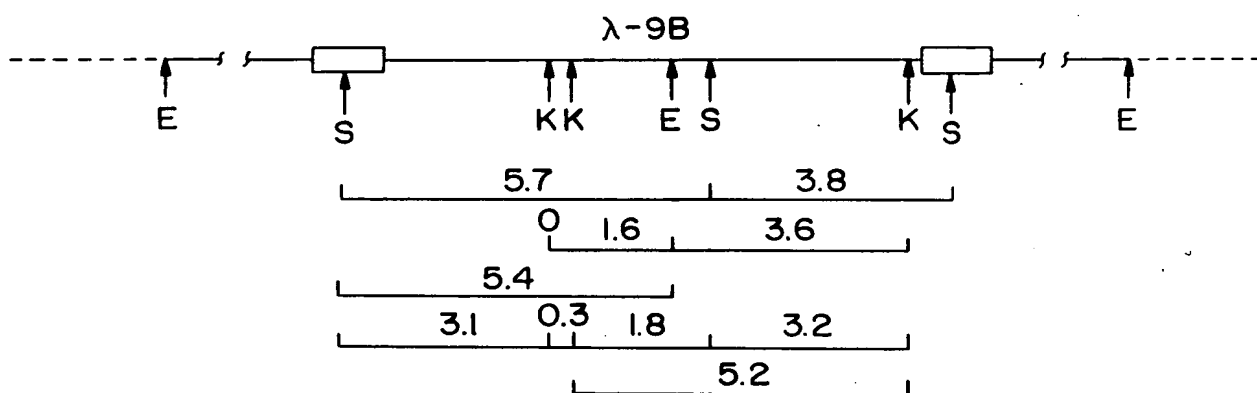


FIG.2

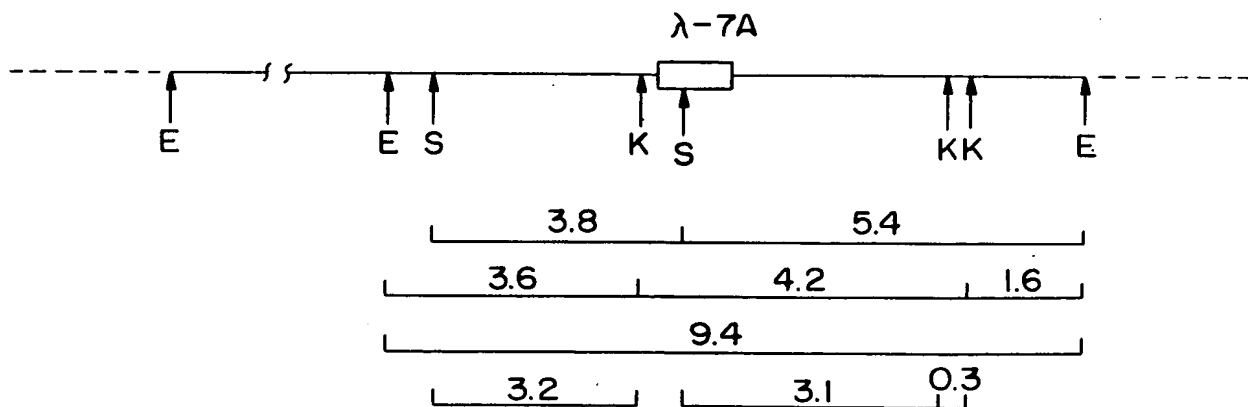


FIG.3

Argument Map in DNA Strand ssarv2
 from the '/v/lib/6mers' file.
 Translation shown at open reading frames.

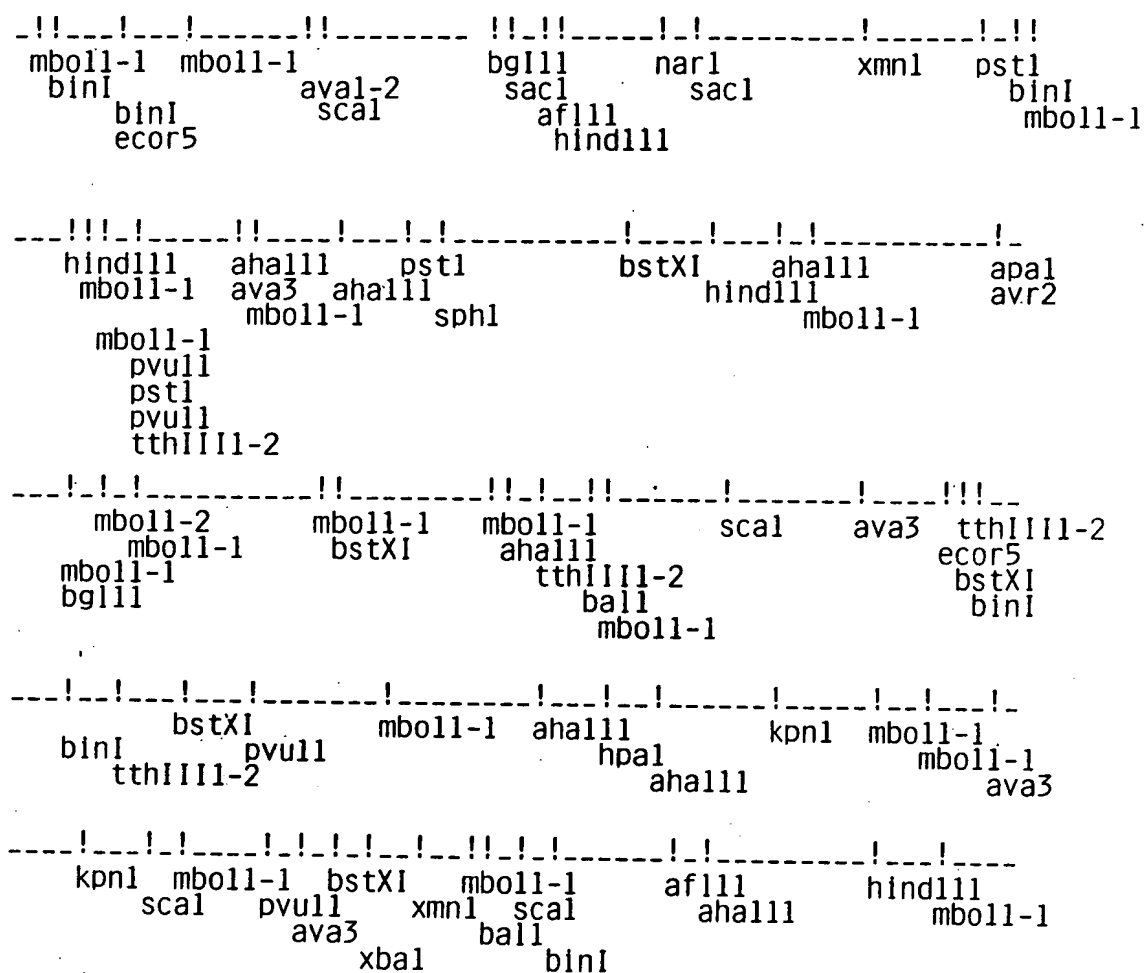


FIG. 4A

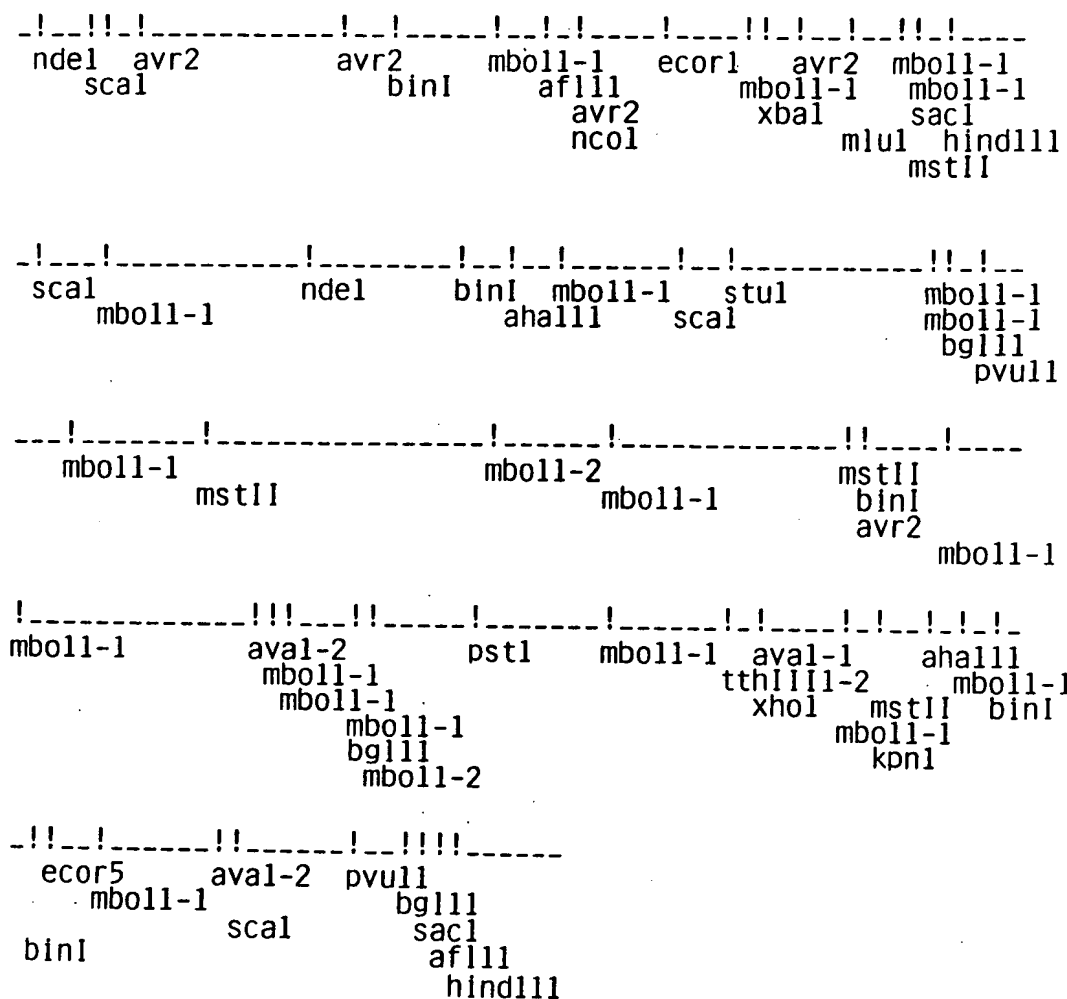


FIG. 4B

- 1 CTGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCACAC
GACCTTCCCGATTAAACCAGGGTTTCTTCTGTTCTCTAGGAAGTACACCTAGATGGTGTG
26 mbo11, 50 bin1,
- 63 ACAAGGCTACTTCCCTGATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCAET
TGTCCGATGAAGGGACTAACCCTTAATGTGTGGTCCCGGTCCCTAGTCTATAGGTGA
107 bin1, 113 ecor5,
- 123 GACCTTTGGATGGTGTCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCAA
CTGGAAACCTACCACGAAGTTCGATCATGGTCAACTCGGTCTCTTCCATCTTCTCCGGTT
172 mbo11,
- 183 TGAAGGAGAGAAACAACAGCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACGCGGA
ACTTCCTCTCTTGTGTCGAACAATGTGGGATACTCGGACGTACCCTACCTCCTGCGCT
243 GAAAGAAGTGTTAGTGTGGAGGTTTGACAGCAAAGTATGATTTTATCATCATGGCCCGAGA
CTTTCTTACAAATCACACCTCCAAAGTGTGTTTATGATCGTAAAGTAGTGTACCGGGCTCT
296 aal,
- 303 GCTGCATCCGGAGTACTACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGC
CGACGTAGGCTCATGATGTTTCTGACGACTGTAGCTCGAAAGATGTTCCCTGAAAGGCG
314 sal,
- 363 TGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATGC
ACCCCTGAAAGGTCCCTCCGCACCGGACCCGCCCTGACCCCTACCGCAGGGAGTCTACG
423 TGCATATAAGCAGACTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAG
ACGTATATTCTGCTGACGAAAAACGGACATGACCCAGAGAGACCAATCTGGTCTAGACTC
474 bgl11,
- 483 CCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCTT
GGACCTCGAGAGACCGATTGATCCCTTGGGTGACGAATTCGGAGTTATTTGAAACGGAA
488 sac1, 518 af111, 532 hind111,
- 543 GAGTGTCTCAAGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCA
CTCACGAAGTTCATCACACACGGGCAGACAACACACTGAGACCATGATCTCTAGGGAGT
603 GACCCTTTTAGTCAGTGTGGAATAATCTCTAGCAGTGGCGCCCGAACAGGGACGCGAAAG
CTGGGAAAAATCAGTCACACCTTTTTAGAGATCGTACCGCGGGCTTGTCCTGCGCTTTC
639 nar1,
- 663 CGAAAGTAGAACCAGAGGAGCTCTCTGACGCGAGGACTCGGCTTGCTGAAGCGCGCACAG
GCTTTCTCTTGGTCTCTCTGAGAGAGCTGCGTCTGAGCCGAACGACTTCGCGCGTGTG
680 sac1,
- 723 CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAATTTTTGACTAGCGGAGGCTAGAAG
GTTCTCCGCTCCCCGCCGCTGACCACTCATGCGTTAAAACTGATCGCTCCGATCTTC
783 GAGAGAGAGATGGGTGCGAGAGCGTCGGTATTAAGCGGGGGAGAATTAGATAAATGGGAA
CTCTCTCTACCCACGCTCTCGCAGCCATAATTCGCCCCCTTAATCTATTTACCTT
- MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGluGAG

FIG. 4C

843 LysIleArgLeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAla
 AAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAGTTAAACATATAGTATGGGCA
 TTTTAAGCCAATTCCGGTCCCCCTTCTTTTTTATATTCAATTTTGTATATCATACCCGT

903 SerArgGluLeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCys
 AGCAGGGAGCTAGAACGATTTCGAGTCAATCCTGGCCTGTTAGAAACATCAGAAGGCTGC
 TCGTCCCTCGATCTTGCTAAGCGTCAGTTAGGACCGGACAATCTTTGTAGTCTTCCGACG

959 pst1,

963 ArgGlnIleLeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSer
 AGACAAATATTGGGACAGCTACAGCCATCCCTTCAGACAGGATCAGAAGAACTTAGATCA
 TCTGTTTATAACCCTGTCGATGTCGGTAGGGAAGTCTGTCTTAGTCTTCTTGAATCTAGT

1002 bin1, 1008 mbol1,

1023 LeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnArgIleAspValLysAspThr
 TTATATAATACAGTAGCAACCCCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACC
 AATATATTATGTCATCGTTGGGAGATAACACATGTAGTTTCTATCTACATTTTCTGTGG

1083 LysGluAlaLeuGluLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGln
 AAGGAAGCTTTAGAGAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAA
 TTCCCTTCGAAATCTCTTCTATCTCCTTCTCGTTTTGTTTTATTCTTTTTCCGTGTCGTT

1087 hind111, 1097 mbol1, 1107 mbol1,

1143 AlaAlaAlaAlaAlaGlyThrGlyAsnSerSerGlnValSerGlnAsnTyrProIleVal
 GCAGCAGCTGCAGCTGGCACAGGAAACAGCAGCCAGGTCAGCCAAAATTACCCTATAGTG
 CGTCGTCGACGTCGACCGTGTCTTTGTCTGTCGGTCCAGTCGGTTTTAATGGGATATCAC

1147 pvu11, 1150 pst1, 1153 pvu11, 1156 tth1111,

1203 GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrp
 CAGAACCTACAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACCTTTAAATGCATGG
 GTCTTGATGTCCCCGTTTACCATGTAGTCCGGTATAGTGGATCTTGAAATTTACGTACC

1250 aha111, 1255 ava3,

1263 ValLysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu
 GTAAAAGTAGTAGAAGAAAAGGCTTTTCAGCCAGAAAGTAATACCCATGTTTTTCAGCATTA
 CATTTTCATCATCTTCTTTTCCGAAAGTCGGGTCTTCATTATGGGTACAAAAGTCGTAAT

1275 mbol1,

1323 SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGln
 TCAGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAA
 AGTCTTCCTCGGTGGGGTGTCTAAATTTGTGGTACGATTTGTGTACCCCCCTGTAGTT

1346 aha111,

1383 AlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal
 GCAGCCATGCAAATGTTAAAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGAGTG
 CGTCGGTACGTTTACAATTTTCTCTGATAGTTACTCTTCGACGTCTTACCCTATCTCAC

1423 pst1,

1443 HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAsp
 CATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAC
 GTAGGTACGTCACGTCCCGGATAACGTGGTCCGGTTTACTCTTGGTTCCCTTCACTG

1451 sph1,

FIG. 4D

1503 IleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro
 ATAGCAGGAACACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT
 TATCGTCCTTGATGATCATGGGAAGTCCTTGTTTATCCTACCTACTGTTTATTAGGTGGA

 1563 IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArg
 ATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGA
 TAGGGTCATCCTCTTTAGATATTTTCTACCTATTAGGACCCTAATTTATTTTATCATTCT

 1623 MetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp
 ATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCTTTAGAGAT
 TACATATCGGGATGGTCGTAAGACCTGTATTCTGTTCTGTTTCTTGGGAAATCTCTA
 1636 bstXI,

 1683 TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsn
 TATGTAGACCGGTTCTATAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAAT
 ATACATCTGGCCAAGATATTTTGAGATTCTCGGCTTGTTTTCGAAGTGTCTACATTTTTTA
 1720 hindIII,

 1743 TrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys
 TGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAGATTGTAAGACTATTTTAAAA
 ACCTACTGTCTTTGGAACAACAGGTTTTACGTTTGGGTCTAACATTCTGATAAAATTTT
 1796 ahaIII,

 1803 AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGly
 GCATTGGGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGA
 CGTAACCCTGGTCGTCGATGTGATCTTCTTTACTACTGTCGTACAGTCCCTCACCCCCCT
 1827 mbolI,

 1863 ProGlyHisLysAlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaAsn
 CCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCCATGAGCCAAGTAACAAATCCAGCTAAC
 GGGCCGGTATTTTCGTTCTCAAACCGACTTCGGTACTCGGTTTCATTGTTTAGGTCGATTG
 p18
 1923 IleMetMetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCys
 ATAATGATGCAGAGAGGCAATTTTAGGAACCAAAGAAAGACTGTAAAGTGTTCATTGT
 TATTACTACGTCTCTCCGTTAAATCCTTGTTTCTTCTGACAATTACAAAGTTAACA
 1983 GlyLysGluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpArg
 GGCAAAGAAGGGCACATAGCCAAAAATTGCAGGGCCCTAGGAAAAAGGGCTGTTGGAGA
 CCGTTTCTTCCCGTGTATCGGTTTTTAACGTCCCGGGGATCCTTTTTCCCGACAACCTCT
 2014 apaI, 2019 avr2,

 2043 CysGlyArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGly
 TGTGGAAGGGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGG
 ACACCTTCCCTTCTGTGGTTTACTTTCTAACGTGACTCTCTGTCCGATTAAAAAATCCC
 2102 mbolI,

 2103 LysIleTrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluPro
 AAGATCTGGCCTTCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCA
 TTCTAGACC6GAAGGATGTTCCCTTCCGGTCCCTTAAAGAAAGTCTCGTCTGGTCTCGGT
 2104 bglIII, 2141 mbolI,

FIG. 4E

2163 ThrAlaProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLys
 ACAGCCCCACCGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAG
 TGTCGGGGTGTTCTTCTCTCGAAGTCCAAACCCCTCCTCTTTTGTGAGGGAGAGTCTTC
 2175 mboll,
 GlnGluProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsn
 2223 CAGGAGCCGATAGACAAGGAAGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAAC
 GTCCTCGGCTATCTGTTCTTGACATAGGAAATTGAAGGGAGTCTAGTGAGAAACCGTTG
 AspProSerSerGlnOC
 2283 GACCCCTCGTCACAATAAGGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGA
 CTGGGGAGCAGTGTATTCTATCCCCCGTTGATTTCCTTCGAGATAATCTATGCTCT
 MetAsnLeuProGlyLysTrpLysProLysMetIle
 2342 GCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATA
 CGTCTACTATGTCATAATCTTCTTTACTTAAACGGTCTTTTACCTTTGGTTTTTACTAT
 2360 mboll, 2375 bstXI,
 GlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIleProValGluIleCys
 2402 GGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGT
 CCCCCTTAACCTCCAAAATAGTTTCACTTGTCTAGTCTATGGACATCTTTAGACA
 GlyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArg
 2462 GGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGA
 CCTGTATTTGATATCCATGTCATAATCATCCTGGATGTGGACAGTTGTATTAACCTTCT
 2517 mboll,
 AsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrVal
 2522 AATCTGTTGACTCAGATTGGTTGTACTTTAAATTTCCCATTAGTCCTATTGAAACTGTA
 TTAGACAACCTGAGTCTAACCAACATGAATTTAAAGGGGTAATCAGGATAACTTTGACAT
 2548 ahal11, 2577 tth1111,
 ProValLysLeuLysProGlyMetAspGlyProLysValLysGlnTrpProLeuThrGlu
 2582 CCAGTAAATTAAGCCAGGAATGGATGGCCCAAAGTTAAGCAATGGCCATTGACAGAA
 GGTCAATTTAATTTTCGGTCCTTACCTACCGGGTTTTCAATTCGTTACCGGTAACCTGTCTT
 2627 ball, 2639 mboll,
 GluLysIleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIleSer
 2642 GAAAAAATAAAGCATTAGTAGAGATATGTACAGAAATGGAAAAGGAAGGGAAAATTTCA
 CTTTTTATTTTCGTAATCATCTCTATACATGTCTTTACCTTTTCCTTCCCTTTTAAAGT
 LysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSer
 2702 AAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGT
 TTTTAACCCGGACTTTTAGGTATGTTATGAGGTCATAAACGATATTTCTTTTTTCTGTCA
 2759 scal,
 ThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrp
 2762 ACTAAATGGAGAAAAGTAGTAGATTTTCAGAGAACTTAATAAAGAACTCAAGACTTCTGG
 TGATTTACCTCTTTTGATCATCTAAAGTCTCTGAATTATTTTCTTGAGTTCTGAAGACC
 GluValGlnLeuGlyIleProHisProGlnGlyOC
 2822 GAAGTTTCAGTTAGGAATACCACACCCGAGGGTTAAAAAAGAAAAAATCAGTAACAGTA
 CTTCAAGTCAATCCTTATGGTGTGGGCGTCCCAATTTTTTCTTTTTTAGTCATTGTCAT

FIG. 4F

2882 TTGGATGTGGGTGATGCATACTTTTCAGTTCCTTAGATAAAGACTTTAGAAAGTATACTG
AACCTACACCCACTACGTATGAAAAGTCAAGGGAATCTATTTCTGAAATCTTTCATATGAC
2895 ava3,

2943 CATTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGG
GTAAATGGTATGGATCATATTTGTTACTCTGTGGTCCCTAATCTATAGTCATGTTACACC
2985 ecor5,

3003 LeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeu
CTGCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTA
GACGGTGTCCCTACCTTTCTAGTGGTCGTTATAAGGTTTCATCGTACTGTTTTTAGAAT
3003 tthIII1, 3006 bstXI, 3021 binI,

3063 GluProPheArgLysGlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyr
GAGCCTTTTAGAAAAACAGAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTAT
CTCGGAAAATCTTTTGTCTTAGGTCTGTATCAATAGATAGTTATGTACCTACTAAACATA

3123 ValGlySerAspLeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHis
GTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAACTGAGACAGCAT
CATCCTAGACTGAATCTTTATCCCGTCGTATCTTGTTTTATCTCCTTGACTCTGTCGTA
3126 binI, 3171 tthIII1,

3183 LeuLeuArgTrpGlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeu
CTGTTGAGGTGGGGATTTACCACACCAGACAAAAACATCAGAAAGAACCTCCATTCCCTT
GACAACTCCACCCCTAAATGGTGTGGTCTGTTTTTGTAGTCTTCTTGAGGTAAGGAA
3234 bstXI,

3243 TrpMetGlyTyrGluLeuHisProAspLysTrpThrValGlnProIleMetLeuProGlu
TGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAATGCTGCCAGAA
ACCTACCCAATACTTGAGGTAGGACTATTTACCTGTCATGTCGGATATTACGACGGTCTT

3303 LysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSer
AAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAAATTGAATTGGGCAAGT
TTTCTGTCGACCTGACAGTTACTGTATGTCTTCAATCACCCCTTTTAACTTAACCCGTTCA
3308 pvuII,

3363 GlnIleTyrAlaGlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyThrLysAla
CAGATTTATGCAGGGATTAAAGTAAAGCAGTTATGTAACTCCTTAGAGGAACCAAAGCA
GTCTAAATACGTCCCTAATTTCAATTCGTCAATACATTTGAGGAATCTCCTTGTTTTCGT

3423 LeuThrGluValIleProLeuThrGluGluAlaGluLeuGluAlaGluAsnArgGlu
CTAACAGAAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAG
GATTGTCTTCATTATGGTGATTGTCTTCTTCGTCTCGATCTTGACCGTCTTTTGTCCCTC
3447 mboII,

3483 IleLeuLysGluProValHisGluValTyrTyrAspProSerLysAspLeuValAlaGlu
ATTCTAAAAGAACCAGTACATGAAGTATATTATGACCCATCAAAAGACTTAGTAGCAGAA
TAAGATTTTCTTGGTCATGTACTTCATATAATACTGGGTAGTTTTCTGAATCATCGTCTT

3543 IleGlnLysGlnGlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluProPheLysAsn
ATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAAT
TATGTCTTCGTCCCGTTCCGGTTACCTGTATAGTTTAAATAGTTCTCGGTAAATTTTAA
3594 ahaIII,

FIG. 4G

3603 LeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeu
CTGAAAACAGGAAAGTATGCAAGGATGAGGGGTGCCACACTAATGATGTAAAACAGTTA
GACTTTTGTCTTTTCATACGTTCTACTCCCCACGGGTGTGATTACTACATTTTGTCAAT
3659 hpa1,

3663 ThrGluAlaValGlnLysValSerThrGluSerIleValIleTrpGlyLysIleProLys
ACAGAGGCAGTGCAAAAAGTATCCACAGAAAGCATAGTAATATGGGGAAAGATTCTAAA
TGTCTCCGTCACGTTTTTCATAGGTGTCTTCGTATCATTATACCCCTTTCTAAGGATTT

3723 PheLysLeuProIleGlnLysGluThrTrpGluAlaTrpTrpMetGluTyrTrpGlnAla
TTTAAACTACCCATACAAAAGGAAACATGGGAAGCATGGTGGATGGAGTATTGGCAAGCT
AAATTTGATGGGTATGTTTTCTTTGTACCCTTCGTACCACCTACCTCATAACCGTTTCTGA
3723 aha111,

3783 ThrTrpIleProGluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGln
ACCTGGATTCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAATTATGGTACCAG
TGGACCTAAGGACTCACCTCAAACAGTTATGGGGAGGGAATCACTTTAATACCATGGTC
3835 kpn1,

3843 LeuGluLysGluProIleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArg
TTAGAGAAAGAACCCATAGTAGGAGCAGAACTTTCTATGTAGATGGGGCAGCTAATAGG
AATCTCTTTCTTGGGTATCATCCTCGTCTTTGAAAGATACATCTACCCCGTCGATTATCC

3903 GluThrLysLeuGlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSer
GAGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGTTGTCTCC
CTCTGATTTAATCCTTTTCGTCTATACAATGACTGTCTCTTCTGTTTTTCAACAGAGG
3943 mbol1,

3963 IleAlaAspThrThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAsp
ATAGCTGACACAACAAATCAGAAGACTGAATTACAAGCAATTCATCTAGCTTTGCAGGAT
TATCGACTGTGTTGTTTAGTCTTCTGACTTAATGTTCTGTTAAGTAGATCGAAACGTCCTA
3983 mbol1,

4023 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAla
TCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCA
AGCCCTAATCTTCATTTGTATCATTGTCTGAGTGTATACGTAATCCTTAGTAAGTTCGT
4060 ava3,

4083 GlnProAspLysSerGluSerGluLeuValSerGlnIleIleGluGlnLeuIleLysLys
CAACCAGATAAGAGTGAATCAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAG
GTTGGTCTATTCTCACTTAGTCTCAATCAGTCAGTTTATTATCTCGTCAATTATTTTTTC

4143 GluLysValTyrLeuAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnVal
GAAAAGGTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTA
CTTTTCCAGATGGACCGTACCCATGGTCGTGTGTTTCCTTAACCTCCTTTACTTGTTTCAT
4163 kpn1,

4203 AspLysLeuValSerAlaGlyIleArgLysValLeuPheLeuAsnGlyIleAspLysAla
GATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTGAATGGAATAGATAAGGCC
CTATTTAATCAGTCACGACCTTAGTCCTTTTCATGATAAAAACCTTACCTTATCTATTCCGG
4232 scal,

FIG. 4H

4263 GlnGluGluHisGluLysTyrHisSerAsnTrpArgAlaMetAlaSerAspPheAsnLeu
CAAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACTG
GTTCTTCTTGTACTCTTTATAGTGTCTTAACCTCTCGTTACCGATCACTAAAATTGGAC
4266 mbol1,

4323 ProProValValAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGlu
CCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAA
GGTGGACATCATCGTTTTCTTTATCATCGGTCGACACTATTTACAGTCGATTTTCTCTT
4352 pvul1,

4383 AlaMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeu
GCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATCTA
CGGTACGTACCTGTTTCATCTGACATCAGGTCCTTATACCGTTGATCTAACATGTGTAGAT
4386 ava3, 4410 bstX1, 4439 xba1,

4443 GluGlyLysIleIleLeuValAlaValHisValAlaSerGlyTyrIleGluAlaGluVal
GAAGGAAAAATTATCCTGGTAGCAGTTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTT
CTTCCTTTTTAATAGGACCATCGTCAAGTACATCGGTCACCTATATATCTTCGTCTTCAA
4497 xmn1,

4503 IleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeuAlaGlyArgTrp
ATTCCAGCAGAGACAGGGCAGGAAACAGCATATTTTCTCTTAAATTTAGCAGGAAGATGG
TAAGGTCGTCTCTGTCCCGTCCTTTGTCGTATAAAAGAGAATTTAATCGTCCTTCTACC
4555 mbol1, 4560 bal1,

4563 ProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerThrThrValLysAla
CCAGTAAAAACAATACATACAGACAATGGCAGCAATTTCCACGACTACTACGGTTAAGGCC
GGTCATTTTTGTTATGTATGTCTGTTACCGTCGTTAAAGTGGTCATGATGCCAATTCGGG
4605 scal,

4623 AlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGln
GCCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAA
CGGACAACCCCGTCCCTAGTTTCGTCTTAAACCGTAAGGGATGTTAGGGGTTTCAGTT
4639 bin1,

4683 GlyValValGluSerMetAsnAsnGluLeuLysLysIleIleGlyGlnValArgAspGln
GGAGTAGTAGAATCTATGAATAATGAATTAAGAAAATTATAGGACAGGTAAGAGATCAG
CCTCATCATCTTAGATACTTATTACTTAATTTCTTTAATATCCTGTCCATTCTCTAGTC
4743 AlaGluHisLeuLysThrAlaValGlnMetAlaValPheIleHisAsnPheLysArgLys
GCTGAACACCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAGAAAA
CGACTTGTGGAATTCTGTCTCATGTTTACCGTCATAAGTAGGTGTTAAATTTTCTTT
4752 alf11, 4791 ahal11,

480 rSerAlaGlyGluArgIleValAspIleIleAlaThrAspIle
CAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATA
GTCACGTCCCCCTTTCTTATCATCTGTATTATCGTTGTCTGTAT

486 qfp12
LysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArg
AAGCAAATTACAAAATTCAAATTTTCGGGTTTTATTACAGG
TTCGTTAATGTTTTTAAGTTTTAAAGCCCAATAATGTCC

FIG. 4 I

AspAsnLysAspProLeuTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAla
 4923 GACAACAAAGATCCCCTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAAGGGGCA
 CTGTTGTTTCTAGGGGAAACCTTTCCTGGTCGTTTCGAAGAGACCTTCCACTTCCCCGT
 4956 hind111,
 ValValIleGlnAspAsnSerAspIleLysValValProArgArgLysAlaLysIleIle
 4983 GTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAAATCATT
 CATCATTATGTTCTATTATCACTGTATTTTCATCACGGTTCTTCTTTTCGTTTTTAGTAA
 5023 mbol1,
 MetGluAsnArgTrpGlnValMetIleValTrpGlnValAspArgMetArgIle
 ArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArgGlnAspGluAsp
 5043 AGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGAT
 TCCCTAATACCTTTTGTCTACCGTCCACTACTAACACACCGTTTCATCTGTCTACTCCTA
 ArgTreTrpLysSerLeuValLysHisHisMetTyrIleSerLysLysAlaLysGlyTrp
 AM
 5103 TAGAACATGGAAAAGTTTAGTAAAACACCATATGTATATTTCAAAGAAAGCTAAAGGATGG
 ATCTTGACCTTTTCAAATCATTTTGTGGTATACATATAAAGTTTCTTTTCGATTTCCTACC
 5131 nde1,
 PheTyrArgHisHisTyrGluSerThrHisProArgValSerSerGluValHisIle
 5163 TTTTATAGACATCACTATGAAAGTACTCATCCAAGAGTAAGTTTCAGAAGTACACATC
 AAAATATCTGTAGTGATACTTTCATGAGTAGGTTCTCATTCAAGTCTTCATGTGTAG
 5185 scal,
 ProLeuGlyAspAlaLysLeuValIleThrThrTyrTrpGlyLeuHisThrGlyGluArg
 5221 CCCCTAGGGGATGCTAAATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGA
 GGGGATCCCCTACGATTTAACCATTATTGTTGTATAACCCAGACGTATGTCCTCTTTCT
 5223 avr2,
 GluTrpHisLeuGlyGlnGlyValAlaIleGluTrpArgLysLysLysTyrSerThrGln
 5281 GAATGGCATTGTTGGGCCAGGGAGTCGCCATAGAATGGAGGAAAAAGAAATATAGCACACAA
 CTTACCGTAAACCCGGTCCCTCAGCGGTATCTTACCTCCTTTTTCTTTATATCGTGTGT
 ValAspProGlyLeuAlaAspGlnLeuIleHisLeuHisTyrPheAspCysPheSerGlu
 5341 GTAGACCCTGGCCTAGCAGACCACTAATTCATCTGCATTATTTTGATTGTTTTTCAGAA
 CATCTGGGACCGGATCGTCTGGTTGATTAAGTAGACGTAATAAACTAACAAAAAGTCTT
 SerAlaIleLysAsnAlaIleLeuGlyTyrArgValSerProArgCysGluTyrGlnAla
 5401 TCTGCTATAAAAAATGCCATATTAGGATATAGAGTTAGTCCTAGGTGTGAATATCAAGCA
 AGACGATATTTTTACGGTATAATCCTATATCTCAATCAGGATCCACACTTATAGTTCGT
 5440 avr2,
 GlyHisAsnLysValGlySerLeuGlnTyrLeuAlaLeuAlaAlaLeuIleThrProLys
 5461 GGACATAACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAAACACCAAAA
 CCTGTATTGTTCCATCTAGAGATGTTATGAACCGTGATCGTCGTAATTATTGTGGTTTT
 5476 bin1,
 LysThrLysProProLeuProSerValLysLysLeuThrGluAspArgTrpAsnLysPro
 5521 AAGACAAAGCCACCTTTGCCTAGTGTTAAGAAACTGACAGAGGATAGATGGAACAAGCCC
 TTCTGTTTCGGTGGAAACGGATCACAATTCTTTGACTGTCTCCTATCTACCTTGTTCGGG

FIG. 4J

- 5581 GlnLysThrLysGlyHisArgGlySerHisThrMetAsnGlyHisAM
CAGAAGACCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGG
GTCTTCTGGTTCCCGGTGTCTCCCTCGGTATGTTACTTACCTGTGATCTCGAAAATCTCC
5583 mbo11,
- 5641 AGCTTAAGAGAGAAGCTGTTAGACATTTTCCTAGGCCATGGCTCCATAGCTTAGGACAAT
TCGAATTCTCTCTTCGACAATCTGTAAAAGGATCCGGTACCGAGGTATCGAATCCTGTTA
5643 afl111, 5670 avr2, 5676 nco1,
- 5701 ATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGGAAGCCATAATAAGAATTCTGC
TATAGATACTTTGAATACCCCTATGAACCCGTCCTCACCTTCGGTATTATTCTTAAGACG
5752 ecor1,
- 5761 AACAACTGCTGTTTATTCATTTTCTTACGAATTGGGTGTCAACATAGCAGAATAGGCATTATTC
TTGTTGACGACAAATAAGTAAAGTCTTAACCCACAGTTGTATCGTCTTATCCGTAATAAG
- 5821 AACAGAGGAGAGCAAGAAGAAATGGAGCCAGTAGATCCTAATCTAGAGCCCTGGAAGCAT
TTGTCTCCTCTCGTTCTTCTTTACCTCGGTCTAGGATTAGATCTCGGGACCTTCGTA
5836 mbo11, 5862 xba1,
- 5881 CCAGGAAGTCAGCCTAGGACTGCTTGTAAACAATTGCTATTGTAAAAAGTGTGCTTTTCAT
GGTCCTTCAGTCGGATCCTGACGAACATTGTTAACGATAACATTTTTTCAACGAAAGTA
5893 avr2,
- 5941 TGCTACGCGTGTTCACAAAGAAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGA
ACGATGCGCACAAAGTGTCTTTTCCGAATCCGTAGAGGATACCGTCCTTCTTCGCTCT
5945 mlu1, 5988 mbo11,
- 6001 CAGCGACGAAGAGCTCCTCAGGACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAA
GTCGCTGCTTCTCGAGGAGTCTGTGAGTCTGAGTAGTTTGAAGAGATAGTTTCGTCATT
6008 mbo11, 6011 sac1, 6016 mstII, 6038 hindIII,
- 6061 GTAGTAAATGTAATGCAATCTTTACAAATATTAGCAATAGTATCATTAGTAGTAGTAGCA
CATCATTTACATTACGTTAGAAATGTTTATAATCGTTATCATAGTAATCATCATCATCGT
- 6121 ATAATAGCAATAGTTGTGTGGACCATAGTACTCATAGAATATAGGAAAATATTAAGACAA
TATTATCGTTATCAACACACCTGGTATCATGAGTATCTTATATCCTTTTATAATTCTGTT
6147 sca1,
- 6181 AGAAAAATAGACAGATTAATTGATAGAATAAGAGAAAAAGCAGAAGACAGTGGCAATGAAA
TCTTTTATCTGTCTAATTAACATCTTATTCTTTTTCTGTTCTGTCAACGTTACTTT
6222 mbo11,
- 6241 ValLysGlyThrArgArgAsnTyrGlnHisLeuTrpArgTrpGlyThrLeuLeuLeuGly
GTGAAGGGGACCAGGAGGAATTATCAGCACTTGTGGAGATGGGGCACCTTGCTCCTTGGG
CACTTCCCCTGGTCTCTTAATAGTCGTGAACACCTCTACCCCGTGAACGAGGAACCC
- 6301 MetLeuMetIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValPro
ATGTTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTACAGTTTATTATGGAGTACCT
TACAACACTAGACATCACGATGTCTTTTAAACACCCAGTGTCAAATAATACCTCATGGA

ENV

FIG. 4K

- 6361 ValTrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAlaArgAlaTyrAspThr
GTGTGGAAAGAAGCAACTACCACTCTATTTTGTGCATCAGATGCTAGAGCATATGATACA
CACACCTTTCTTCGTTGATGGTGAGATAAAACACGTAGTCTACGATCTCGTATACTATGT
6410 ndel,
- 6421 GluValHisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGlu
GAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAA
CTCCATGTATTACAAACCCGGTGTGTACGGACACATGGGTGTCTGGGGTTGGGTGTTCTT
- 6481 ValValLeuGlyAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGln
GTAGTATTGGGAAATGTGACAGAAAATTTTAAACATGTGGAAAAATAACATGGTAGAACAG
CATCATAACCCCTTTACACTGTCTTTTAAATTGTACACCTTTTTATTGTACCATCTTGTC
- 6541 MetGlnGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysValLysLeuThr
ATGCAGGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACC
TACGTCCTCCTATATTAGTCAAATACCTAGTTTCGGATTTCGGTACACATTTTAATTGG
6567 binl,
- 6601 ProLeuCysValThrLeuAsnCysThrAspLeuGlyLysAlaThrAsnThrAsnSerSer
CCACTCTGTGTTACTTTAAATTGCACTGATTTGGGGAAGGCTACTAATACCAATAGTAGT
GGTGAGACACAATGAAATTTAACGTGACTAAACCCCTTCCGATGATTATGGTTATCATCA
6615 ahal11,
- 6661 AsnTrpLysGluGluIleLysGlyGluIleLysAsnCysSerPheAsnIleThrThrSer
AATTGGAAAGAAGAAATAAAAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGC
TTAACCTTTCTTTTATTTTCTCTTTATTTTTTGACGAGAAAAGTTATAGTGGTGTTCG
6670 mbol1,
- 6721 IleArgAspLysIleGlnLysGluAsnAlaLeuPheArgAsnLeuAspValValProIle
ATAAGAGATAAGATTTCAGAAAGAAAATGCACCTTTTTTCGTAACCTTGATGTAGTACCAATA
TATTCTCTATTCTAAGTCTTTCTTTTACGTGAAAAGCATTGGAACATCATGTTAT
- 6781 AspAsnAlaSerThrThrThrAsnTyrThrAsnTyrArgLeuIleHisCysAsnArgSer
GATAATGCTAGTACTACTACCAACTATACCAACTATAGGTTGATACATTGTAACAGATCA
CTATTACGATCATGATGATGGTTGATATGGTTGATATCCAACATGTAAACATTGTCTAGT
6790 scal,
- 6841 ValIleThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCysThr
GTCATTACACAGGCCTGTCCAAAGGTATCATTTGAGCCAATTCCCATACATTATTGTACC
CAGTAATGTGTCCGGACAGGTTTCCATAGTAAACTCGGTTAAGGGTATGTAATAACATGG
6851 stul,
- 6901 ProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyLysGlyProCys
CCGGCTGGTTTTGCGATTCTAAAGTGTAATAATAAACGTTCAATGGAAAAGGACCATGT
GGCCGACCAAAACGCTAAGATTTACATTATTATTTGCAAGTTACCTTTTCTGGTACA
- 6961 ThrAsnValSerThrValGlnCysThrHisGlyIleArgProIleValSerThrGlnLeu
ACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAATAGTGTCAACTCAACTG
TGTTTACAGTCGTGTCATGTTACATGTGTACCTTAATCCGGTTATCACAGTTGAGTTGAC
- 7021 LeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerAspAsnPheThrAsn
CTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTACGAAC
GACAATTTACCGTCAGATCGTCTTCTTCTCCATCATTAATCTAGACTGTTAAAGTGCTTG
7042 mbol1, 7045 mbol1, 7060 bgl11,

FIG. 4L

7081 AsnAlaLysThrIleIleValGlnLeuAsnGluSerValAlaIleAsnCysThrArgPro
 AATGCTAAAACCATAGTACAGCTGAATGAATCTGTAGCAATTAAGTGTACAAGACCC
 TTACGATTTTGGTATTATCATGTCGACTTACTTAGACATCGTTAATTGACATGTTCTGGG
 7102 pvu11,
 7141 AsnAsnAsnThrArgLysSerIleTyrIleGlyProGlyArgAlaPheHisThrThrGly
 AACAAACAATACAAGAAAAAGTATCTATATAGGACCAGGGAGAGCATTTCATACAACAGGA
 TTGTTGTTATGTTCTTTTTCATAGATATATCCTGGTCCCTCTCGTAAAGTATGTTGTCCT
 7199 mbo11,
 7201 ArgIleIleGlyAspIleArgLysAlaHisCysAsnIleSerArgAlaGlnTrpAsnAsn
 AGAATAATAGGAGATATAAGAAAAGCACATTGTAACATTAGTAGAGCACAATGGAATAAC
 TCTTATTATCCTCTATATTCTTTTCGTGTAACATTGTAATCATCTCGTGTTACCTTATTG
 7261 ThrLeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsnAsnLysThrIleVal
 ACTTTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGGAATAATAAAACAATAGTC
 TGAAATCTTGCTATCAATTTTTTAATTCTCTTGTCAAACCTTATTATTTTGTATCAG
 7321 PheAsnGlnSerSerGlyGlyAspProGluIleValMetHisSerPheAsnCysArgGly
 TTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTAAATTGTAGAGGG
 AAATTAGTTAGGAGTCTCCCTGGGTCTTTAACATTACGTGTCAAATTAACATCTCCC
 7331 mstII,
 7381 GluPhePheTyrCysAsnThrThrGlnLeuPheAsnAsnThrTrpArgLeuAsnHisThr
 GAATTTTCTACTGTAATACAACAACACTGTTTAATAATACATGGAGGTTAAATCACACT
 CTTAAAAAGATGACATTATGTTGTGTTGACAAATTATTATGTACCTCCAATTTAGTGTGA
 7441 GluGlyThrLysGlyAsnAspThrIleIleLeuProCysArgIleLysGlnIleIleAsn
 GAAGGAACTAAAGGAAATGACACAATCATACTCCCATGTAGAATAAAACAAATTATAAAC
 CTTCTTGATTTCCTTTACTGTGTTAGTATGAGGGTACATCTTATTTTGTTAATATTTG
 7501 MetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGlyGlyGlnIleSerCys
 ATGTGGCAGGAAGTAGGAAAAAGCAATGTATGCCCTCCCATTTGGAGGACAAATTAGTTGT
 TACACCGTCCTTCATCCTTTTCTGTACATACGGGGAGGGTAACCTCCTGTTTAATCAACA
 7561 SerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThrAsnValThrAsnAsp
 TCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTACAAATGTAACATAATGAC
 AGTAGTTTATAATGTCCCACGATAATTGTTCTCTACCACCATGTTTACATTGATTACTG
 7621 ThrGluValPheArgProGlyGlyGlyAspMetArgAspAsnTrpArgSerGluLeuTyr
 ACCGAGGTCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATAT
 TGGCTCCAGAAGTCTGGACCTCCTCCTATACTCCCTGTTAACCTCTTCACTTAATATA
 7628 mbo11,
 7681 LysTyrLysValIleLysIleGluProLeuGlyIleAlaProThrLysAlaLysArgArg
 AAATATAAAGTAATAAAAAATTGAACCATTAGGAATAGCACCCACCAAGGCAAAGAGAAGA
 TTTATATTTTCAATTATTTTAACTTGGAATCCTTATCGTGGGTGGTCCGTTTCTCTCT
 7736 mbo11,
 7741 ValValGlnArgGluLysArgAlaValGlyIleValGlyAlaMetPheLeuGlyPheLeu
 GTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGTAGGAGCTATGTTCTTGGGTTCTTG
 CACCACGTCTCTTTTTTCTCGTCACCCCTTATCATCTCGATACAAGGAACCCAAGAAC
 7801 GlyAlaAlaGlySerThrMetGlyAlaValSerLeuThrLeuThrValGlnAlaArgGln
 GGAGCAGCAGGAAGCACTATGGGCGCAGTGTCTATTGACGCTGACGGTACAGGCCAGACAA
 CCTCGTCGTCTTCGTGATACCCGCGTCACAGTAACTGCGACTGCCATGTCCGGTCTGTT

FIG. 4M

7861 LeuLeuSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGln
 TTATTGCTGGTATAGTGCAACAGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAACAA
 AATAACAGACCATATCACGTTGTCGTTCTGTTAAACGACTCCCGATAACTCCGCGTTGTT
 7921 HisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaVal
 CATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTGTG
 GTAGACAACGTTGAGTGTGAGACCCCGTAGTTCTGTCGAGGTCCGTTCTCAGGACCGACAC
 7981 GluArgTyrLeuArgAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIle
 GAAAGATACCTAAGGGATCAACAGCTCCTAGGGATTTGGGGTTGCTCTGGAAAACCTCATT
 CTTTCTATGGATTCCCTAGTTGTGAGGATCCCTAAACCCCAACGAGACCTTTTGAGTAA
 7989 mstII, 7995 binI, 8007 avr2,
 8041 CysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluAspIleTrp
 TGCACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAAGACATTTGG
 ACGTGGTGACGACACGGAACCTTACGATCAACCTCATTATTTAGAGACCTTCTGTAAACC
 8089 mbolI,
 8101 AspAsnMetThrTrpMetGlnTrpGluArgGluIleAspAsnTyrThrAsnThrIleTyr
 GATAACATGACCTGGATGCAGTGGGAAAGAGAAATTTGACAATTACACAAACACAATATAC
 CTATTGTACTGGACCTACGTCACCCCTTCTCTTTAACTGTTAATGTGTTTGTGTTATATG
 8161 ThrLeuLeuGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu
 ACCTTACTTGAAGAATCGCAGAACCAACAGAAAGAAATGAACAAGAATTATTAGAATTG
 TGAATGAACCTTCTTAGCGTCTTGGTTGTTCTTTCTTACTTGTCTTAATAATCTTAAC
 8170 mbolI,
 8221 AspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrAsnTrpLeuTrpTyrIleLys
 GATAAGTGGGCAAGTTTGTGGAATTGGTTTAGCATAACAACTGGCTGTGGTATATAAAG
 CTATTCACCCGTTCAAACACCTTAACCAAATCGTATTGTTTGACCGACACCATATATTTCT
 8281 IlePheIleMetIleValGlyGlyLeuValGlyLeuArgIleValPheAlaValLeuSer
 ATATTGATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTGTGCTGTGCTTTCT
 TATAAGTATTACTATCATCCTCCGAACCATCAAATTCTTATCAAAAACGACACGAAAGA
 8341 IleValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrArgLeuProVal
 ATAGTGAATAGAGTTAGGCAGGGATACTCACCATTGTCATTTTCAGACCCGCCTCCAGTC
 TATCACTTATCTCAATCCGTCCCTATGAGTGGTAACAGTAAAGTCTGGGCGGAGGGTCAG
 8400 aval,
 8401 ProArgGlyProAspArgProAspGlyIleGluGluGluGlyGlyGluArgAspArgAsp
 CCGAGGGGACCCGACAGGCCCGACGGAATCGAAGAAGAAGGTGGAGAGAGACAGAGAC
 GGCTCCCCTGGGCTGTCCGGGCTGCCTTAGCTTCTTCTTCCACCTCTCTCTGTCTCTG
 8431 mbolI, 8434 mbolI,
 8461 ArgSerValArgLeuValAspGlyPheLeuAlaLeuIleTrpGluAspLeuArgSerLeu
 AGATCCGTTTCGATTAGTGGATGGATTCTTAGCACTTATCTGGGAAGATCTGCGGAGCCTG
 TCTAGGCAAGCTAATCACCTACCTAAGAATCGTGAATAGACCTTCTTAGACGCCTCGGAC
 8503 mbolI, 8505 bgII,
 8521 CysLeuPheSerTyrArgArgLeuArgAspLeuLeuLeuIleAlaAlaArgThrValGlu
 TGCCTCTTCAGCTACCGCCGCTTGAGAGACTTACTCTTGATTGCAGCGAGGACTGTGGAA
 ACGGAGAAGTCGATGGCGGCGAACTCTCTGAATGAGAACTAACGTCGCTCCTGACACCTT
 8525 mbolI,

FIG. 4N

8581 IleLeuGlyHisArgGlyTrpGluAlaLeuLysTyrTrpTrpSerLeuLeuGlnTyrTrp
 ATTCTGGGGCACAGGGGGTGGGAAGCCCTCAAATATTGGTGGAGTCTCCTGCAGTATTGG
 TAAGACCCCGTGTCCCCACCCCTTCGGGAGTTTATAACCACTCAGAGGACGTCATAACC

8629 pst1,

8641 IleGlnGluLeuLysAsnSerAlaValSerTrpLeuAsnAlaThrAlaIleAlaValThr
 ATTCAGGAACATAAGAATAGTGTCTGTTAGCTGGCTCAACGCCACAGCTATAGCAGTAAC
 TAAGTCCTTGATTCTTATCACGACAATCGACCGAGTTGCGGTGTCGATATCGTCATTGA

8701 GluGlyThrAspArgValIleGluValAlaGlnArgAlaTyrArgAlaIleLeuHisIle
 GAGGGGACAGATAGGGTTATAGAAGTAGCACAAAGAGCTTATAGAGCTATTCTCCACATA
 CTCCCCTGTCTATCCCAATATCTTCATCGTGTCTCTCGAATATCTCGATAAGAGGTGTAT

8761 HisArgArgIleArgGlnGlyLeuGluArgLeuLeuLeuOC MetGlyGlyLysTrpSer
 CATAGAAGAATTAGACAGGGCTTGGAAAGGCTTTTGCTATAAGATGGGTGGCAAGTGGTCA
 GTATCTTCTTAATCTGTCCCGAACCTTTCGAAAACGATATTCTACCCACCGTTCACCACT

8765 mbol1,

8822 LysArgSerMetGlyGlyTrpSerAlaIleArgGluArgMetArgArgAlaGluProArg
 AAACGTAGTATGGGTGGATGGTCTGCTATAAGGGAAAGAATGAGACGAGCTGAGCCACGA
 TTTGCATCATACCCACCTACCAGACGATATTCCCTTTCTTACTCTGCTCGACTCGGTGCT

8882 AlaGluProAlaAlaAspGlyValGlyAlaValSerArgAspLeuGluLysHisGlyAla
 GCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTGGAAAAACATGGAGCA
 CGACTCGGTCGTCGTCTACCCACCCCTCGTCATAGAGCTCTGGACCTTTTGTACCTCGT

8883 tth1111, 8916 aval xho1,

8942 IleThrSerSerAsnThrAlaAlaThrAsnAlaAspCysAlaTrpLeuGluAlaGlnGlu
 ATCACAAGTAGCAATACAGCAGCTACTAATGCTGATTGTGCCTGGCTAGAAGCACAAGAG
 TAGTGTTTCATCGTTATGTCGTCGATGATTACGACTAACACGGACCGATCTTCGTGTTCTC

9002 GluGluGluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLys
 GAGGAAGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGACTTACAAG
 CTCCTTCTCCACCCAAAAGGTACGTCTGGAGTCCATGGAAATTCTGGTTACTGAATGTTCT

9005 mbol1, 9029 mstII, 9034 kpn1,

9062 AlaAlaLeuAspIleSerHisPheLeuLysGluLysGlyGlyLeuGluGlyLeuIleTrp
 GCAGCTTTAGATATTAGCCACTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTTGG
 CGTCGAAATCTATAATCGGTGAAAATTTTCTTTTCCCCCTGACCTTCCCGATTAAACC

9085 aha111,

9122 SerGlnArgArgGlnGluIleLeuAspLeuTrpIleTyrHisThrGlnGlyTyrPhePro
 TCCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCT
 AGGGTTTCTTCTGTTCTCTAGGAACCTAGACACCTAGATGGTGTGTGTTCCGATGAAGGGA

9129 mbol1, 9153 bnl1,

9182 AspTrpGlnAsnTyrThrProGlyProGlyIleArgTyrProLeuThrPheGlyTrpCys
 GATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGC
 CTAACCGTCTTAATGTGTGGTCCCGGTCCCTAGTCTATAGGTGACTGGAAACCTACCACG

9210 bnl1, 9216 ecor5,

FIG. 40

9242 PheLysLeuValProValGluProGluLysValGluGluAlaAsnGluGlyGluAsnAsn
TTCAAGCTAGTACCAAGTTGAGCCAGAGAAGGTAGAAGAGGCCAATGAAGGAGAGAACAAC
AAGTTCGATCATGGTCAACTCGGTCTCTTCCATCTTCTCCGGTTACTTCCTCTCTTGTG
9275 mbo11,

9302 SerLeuLeuHisProMetSerLeuHisGlyMetGluAspAlaGluLysGluValLeuVal
AGCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACGCGGAGAAAGAAAGTGTAGTG
TCGAACAATGTGGGATACTCGGACGTACCCTACCTCCTGCGCCTCTTCTTCACAATCAC

9362 TrpArgPheAspSerLysLeuAlaPheHisHisMetAlaArgGluLeuHisProGluTyr
TGGAGGTTTGACAGCAAAGTACGATTTTCATCACATGGCCCGAGAGCTGCATCCGGAGTAC
ACCTCCAAAGTGTCTGTTGATCGTAAAGTAGTGTACCGGCTCTCGACGTAGGCCCTCATG
9399 aal, 9417 sac1,

9422 TyrLysAspCysOP
TACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGG
ATGTTTCTGACGACTGTAGCTCGAAAGATGTTCCCTGAAAGGCGACCCCTGAAAGGTCCC
9482 AGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATGCTGCATATAAGCAGCTG
TCCGCACCGGACCCGCCCTGACCCCTCACCGCAGGGAGTCTACGACGTATATTCGTCGAC
9536 pvu11,

9542 CTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC
GAAAAACGGACATGACCCAGAGAGACCAATCTGGTCTAGACTCGGACCCTCGAGAGACCG
9576 bgl111, 9590 sac1,

9602 TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTG
ATTGATCCCTTGGGTGACGAATTCGGAGTTATTTTGAACGGAACCTCACGAAGTTTCATCAC
9620 af111, 9634 hind111,

9662 TGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTG
ACACGGGCAGACAACACACTGAGACCATTGATCTCTAGGGAGTCTGGGAAAATCAGTCAC

9722 TGGAAAAATCTCTAGCAG
ACCTTTTCTAGAGATCGTC

FIG. 4P

U3 →
 -453 CTGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCAC
 ACACAAGGCTACTTCCCTGATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCA
 -333 CTGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCC
 AATGAAGGAGAGAACAACAGCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACGCG L
 -214 GAGAAAGAAGTGTTAGTGTGGAGGTTTGCAGCAAACTAGCATTTTCATCACATGGCCCCG
 GAGCTGCATCCGGAGTACTACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCG T
 -93 CTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATG
 CTGCATATAAGCAGCTGCTTTTTGCCTGTACTG ← U3 R → GGTCTCTCTGGTTAGACCAGATCTGAG R
 28 CCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTT
 ← R U5 →
 GAGTGCTTCA AGTAGTGTGTGCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCA
 ← U5
 148 GACCCTTTTAGTCAGTGTGGAAAAATCTCTAGCAG TGGCGCCCGAACAGGGACGCGAAA
 GCGAAAGTAGAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACAG
 268 CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAATTTTTGACTAGCGGAGGCTAGAAG
 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGlu 17
 GAGAGAGAGATGGGTGCGAGAGCGTCGGTATTAAGCGGGGAGAAATTAGATAAATGGGAA
 LysIleArgLeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAla
 388 AAAATTCGGTTAAGGCCAGGGGAAAGAAAAAATATAAGTTAAACATATAGTATGGGCA
 SerArgGluLeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCys 57
 AGCAGGGAGCTAGAACGATTTCGAGTCAATCCTGGCCTGTTAGAAACATCAGAAGGCTGC
 ArgGlnIleLeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSer
 508 AGACAAATATTGGGACAGCTACAGCCATCCCTTCAGACAGGATCAGAAGAACTTAGATCA
 LeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnArgIleAspValLysAspThr 97
 TTATATAATACAGTAGCAACCCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACC
 LysGluAlaLeuGluLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGln
 628 AAGGAAGCTTTAGAGAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAA
 AlaAlaAlaAlaAlaGlyThrGlyAsnSerSerGlnValSerGlnAsnTyrProIleVal 137
 GCAGCAGCTGCAGCTGGCACAGGAAACAGCAGCCAGGTCAGCCAAAATTACCCTATAGTG
 GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrp
 748 CAGAACCTACAGGGGCAAATGGTACATCAGGCCATATCACCTAGAAGTTTAAATGCATGG
 ValLysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu 177
 GTAAAAGTAGTAGAAGAAAAGGCTTTTCAGCCCAGAAGTAATACCCATGTTTTTCAGCATTA
 SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGln
 868 TCAGAAGGAGCCACCCCAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAA

FIG. 5A

AlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal 217 G
 GCAGCCATGCAAATGTAAAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGAGTG
 988 HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAsp
 CATCCAGTGCATGCAGGGCCTATTGCACCAAGGCCAAATGAGAGAACCAAGGGGAAGTGAC A
 IleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro 257
 ATAGCAGGAACCTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT G
 1108 IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArg
 ATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGA
 MetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp 297
 ATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAGGAACCCCTTTAGAGAT
 1228 TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsn
 TATGTAGACCGTTCTATAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAAT
 TrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys 337
 TGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAGATTGTAAGACTATTTTAAAA
 1348 AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGly
 GCATTGGGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGA
 ProGlyHisLysAlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaAsn 377
 CCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCCATGAGCCAAGTAACAAATCCAGCTAAC
 1468 IleMetMetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCys
 ATAATGATGCAGAGAGGCAATTTTAGGAACCAAGAAAGACTGTTAAGTGTTTCAATTGT
 GlyLysGluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpArg 417
 GGCAAGAAGGGCACATAGCCAAAATTTGACAGGGCCCTAGGAAAAAGGGCTGTTGGAGA
 CysGlyArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGly
 1588 TGTGGAAGGGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGG
 PhePheArgG
 LysIleTrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluPro 457
 luAspLeuAlaPheLeuGlnGlyLysAlaArgGluPheSerSerGluGlnThrArgAla 23
 AAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCA
 ThrAlaProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLys
 1708 AsnSerProThrArgArgGluLeuGlnValTrpGlyGlyGluAsnAsnSerLeuSerGluA
 ACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACCTCCCTCTCAGAAG P
 GlnGluProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsn 497
 laGlyAlaAspArgGlnGlyThrValSerPheAsnPheProGlnIleThrLeuTrpGln 63
 CAGGAGCCGATAGACAAGGAAGTGTATCCTTTAACTTCCCTCAGATCACTCTTGGCAAC O
 AspProSerSerGlnOC
 ArgProLeuValThrIleArgIleGlyGlyGlnLeuLysGluAlaLeuLeuAspThrGlyA
 1828 GACCCCTCGTCACAATAAGGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAG L
 laAspAspThrValLeuGluGluMetAsnLeuProGlyLysTrpLysProLysMetIle 103
 CAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAG
 GlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIleProValGluIleCysG
 1948 GGGGAATTGGAGGTTTTATCAAAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTG

FIG. 5B

lyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArg 143
 GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA
 AsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValP
 2068 ATCTGTTGACTCAGATTGGTTGTACTTTAAATTTCCCATAGTCCTATTGAAACTGTAC
 roValLysLeuLysProGlyMetAspGlyProLysValLysGlnTrpProLeuThrGlu 183
 CAGTAAATTAAGCCAGGAATGGATGGCCAAAAGTTAAGCAATGGCCATTGACAGAAG
 GluLysIleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIleSerL
 2188 .AAAAAATAAAGCATTAGTAGAGATATGTACAGAAATGGAAAAGGAAGGGAAAATTTCAA
 ysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSer 223
 AAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGTA
 ThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrpG
 2308 CTAATGGAGAAAACTAGTAGATTTTCAGAGAACTTAATAAAGAAGCTCAAGACTTCTGGG
 luValGlnLeuGlyIleProHisProAlaGlyLeuLysLysLysLysSerValThrVal 263
 AAGTTCAGTTAGGAATACCACACCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTAT
 LeuAspValGlyAspAlaTyrPheSerValProLeuAspLysAspPheArgLysTyrThrA
 2428 TGGATGTGGGTGATGCATACTTTTCAGTTCCTTAGATAAAGACTTTAGAAAGTATACTG
 laPheThrIleProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnVal 303
 CATTACCATACCTAGTATAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGC
 LeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuG
 2548 TGCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAG
 luProPheArgLysGlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyr 343
 AGCCTTTTAGAAAACAGAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATG
 ValGlySerAspLeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHisL
 2668 TAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAAGCTGAGACAGCATC
 euLeuArgTrpGlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeu 383
 GTTGAGGTGGGGATTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCTTT
 TrpMetGlyTyrGluLeuHisProAspLysTrpThrValGlnProIleMetLeuProGluL
 2788 GGATGGGTTATGAAGTCCATCCTGATAAATGGACAGTACAGCCTATAATGCTGCCAGAAA
 ysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSer 423
 AAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAAATTGAATTGGGCAAGTC
 GlnIleTyrAlaGlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyThrLysAlaL
 2908 AGATTTATGCAGGGATTAAAGTAAAGCAGTTATGTAAACTCCTTAGAGGAACCAAAGCAC
 euThrGluValIleProLeuThrGluGluAlaGluLeuGluLeuAlaGluAsnArgGlu 463 P
 TAACAGAAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAGA
 IleLeuLysGluProValHisGluValTyrTyrAspProSerLysAspLeuValAlaGluI
 3028 TTCTAAAAGAACCAGTACATGAAGTATATTATGACCCATCAAAGACTTAGTAGCAGAAA
 leGlnLysGlnGlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluProPheLysAsn 503 O
 TACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATC
 LeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuT
 3148 TGAAAACAGGAAAGTATGCAAGGATGAGGGGTGCCACACTAATGATGTAAAACAGTTAA
 hrGluAlaValGlnLysValSerThrGluSerIleValIleTrpGlyLysIleProLys 543 L
 CAGAGGCAGTGCAAAAAGTATCCACAGAAAGCATAGTAATATGGGGAAAGATTCTAAAT

FIG. 5C

PheLysLeuProIleGlnLysGluThrTrpGluAlaTrpTrpMetGluTyrTrpGlnAlaT
 3268 TTAACCTACCCATACAAAAGGAAACATGGGAAGCATGGTGGATGGAGTATTGGCAAGCTA
 hrTrpIleProGluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGln 583
 CCTGGATTCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAATTATGGTACCAGT
 LeuGluLysGluProIleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArgG
 3388 TAGAGAAAGAACCCATAGTAGGAGCAGAACTTTCTATGTAGATGGGGCAGCTAATAGGG
 luThrLysLeuGlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSer 623
 AGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAAGTTGTCTCCA
 IleAlaAspThrThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAspS
 3508 TAGCTGACACAACAAATCAGAAAGACTGAATTACAAGCAATTCATCTAGCTTTGCAGGATT
 erGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAla 663
 CGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCAC
 GlnProAspLysSerGluSerGluLeuValSerGlnIleIleGluGlnLeuIleLysLysG
 3628 AACCAGATAAGAGTGAATCAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGG
 luLysValTyrLeuAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnVal 703
 AAAAGGTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAG
 AspLysLeuValSerAlaGlyIleArgLysValLeuPheLeuAsnGlyIleAspLysAlaG
 3748 ATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTGAATGGAATAGATAAGGCC
 InGluGluHisGluLysTyrHisSerAsnTrpArgAlaMetAlaSerAspPheAsnLeu 743
 AAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGC
 ProProValValAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluA
 3868 CACCTGTAGTAGCAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAG
 laMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeu 783
 CCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATCTAG
 GluGlyLysIleIleLeuValAlaValHisValAlaSerGlyTyrIleGluAlaGluValI
 3988 AAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTA
 leProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeuAlaGlyArgTrp 823
 TTCCAGCAGAGACAGGGCAGGAAACAGCATATTTTCTCTTAAATAGCAGGAAGATGGC
 ProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerThrThrValLysAlaA
 4108 CAGTAAAAACAATACATACAGACAATGGCAGCAATTCACCAGTACTACGGTTAAGGCCG
 laCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGln 863
 CCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAG
 GlyValValGluSerMetAsnAsnGluLeuLysLysIleIleGlyGlnValArgAspGlnA
 4228 GAGTAGTAGAATCTATGAATAATGAATTAAAGAAAAATTATAGGACAGGTAAGAGATCAGG
 laGluHisLeuLysThrAlaValGlnMetAlaValPheIleHisAsnPheLysArgLys 903
 CTGAACACCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAGAAAAAG
 GlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleIleAlaThrAspIleG
 4348 GGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATAC
 InThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArg 943
 AAACCTAAAGAACTACAAAAGCAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGG

FIG. 5D

AspAsnLysAspProLeuTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAlaV
 4468 ACAACAAAGATCCCCTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAAGGGGCAG
 alValIleGlnAspAsnSerAspIleLysValValProArgArgLysAlaLysIleIle 983
 TAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAAATCATT
 ArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArgGlnAspGluAspA
 4588 GGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATT
 M
 AGAACATGGAAAAGTTTAGTAAACACCATATGTATATTTCAAAGAAAGCTAAAGGATGG
 4708 TTTTATAGACATCACTATGAAAGTACTCATCCAAGAGTAAGTTCAGAAGTACACATCCCC
 CTAGGGGATGCTAAATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGAA
 4828 TGGCATTGTTGGGCCAGGGAGTCGCCATAGAATGGAGGAAAAAGAAATATAGCACACAAGTA
 GACCCTGGCCTAGCAGACCAACTAATTCATCTGCATTATTTTGATTGTTTTTCAGAATCT
 4948 GCTATAAAAAATGCCATATTAGGATATAGAGTTAGTCCTAGGTGTGAATATCAAGCAGGA
 CATAACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTATAACACCAAAAAAG
 5068 ACAAAGCCACCTTTGCCTAGTGTTAAGAACTGACAGAGGATAGATGGAACAAGCCCCAG
 AAGACCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAGC
 5188 TTAAGAGAGAAGCTGTTAGACATTTTCTAGGCCATGGCTCCATAGCTTAGGACAATATA
 TCTATGAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAAC
 5308 AACTGCTGTTTATTCATTTTCAAGATTGGGTGTCAACATAGCAGAATAGGCATTATTCAAC
 AGAGGAGAGCAAGAAGAAATGGAGCCAGTAGATCCTAATCTAGAGCCCTGGAAGCATCCA
 5428 GGAAGTCAGCCTAGGACTGCTTGTAAACAATTGCTATTGTAAAAAGTGTGCTTTTATTGC
 TACGCGTGTTTCACAAGAAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAG
 5548 CGACGAAGAGCTCCTCAGGACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAAGTA
 GTAAATGTAATGCAATCTTTACAAATATTAGCAATAGTATCATTAGTAGTAGTAGCAATA
 5668 ATAGCAATAGTTGTGTGGACCATAGTACTCATAGAATATAGGAAAATATTAAGACAAAGA
 AAATAGACAGATTAATTGATAGAATAAGAGAAAAAGCAGAAGACAGTGGCAATGAAAGTG MetLysVal 3
 LysGlyThrArgArgAsnTyrGlnHisLeuTrpArgTrpGlyThrLeuLeuLeuGlyMet
 5788 AAGGGGACCAGGAGGAATTATCAGCACTTGTGGAGATGGGGCACCTTGCTCCTTGGGATG
 LeuMetIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProVal 43
 TTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTACAGTTTATTATGGAGTACCTGTG
 TrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAlaArgAlaTyrAspThrGlu
 5908 TGGAAAGAAGCAACTACCACTCTATTTTGTGCATCAGATGCTAGAGCATATGATACAGAG
 ValHisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGluVal 83
 GTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTA

FIG. 5E

6028 ValLeuGlyAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGlnMet
 GTATTGGGAAATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAGATG
 GlnGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysValLysLeuThrPro 123
 CAGGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCA
 6148 LeuCysValThrLeuAsnCysThrAspLeuGlyLysAlaThrAsnThrAsnSerSerAsn
 CTCTGTGTTACTTTAAATTGCACTGATTTGGGGAAGGCTACTAATACCAATAGTAGTAAT
 TrpLysGluGluIleLysGlyGluIleLysAsnCysSerPheAsnIleThrThrSerIle 163
 TGGAAAGAAGAAATAAAAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGCATA
 6268 ArgAspLysIleGlnLysGluAsnAlaLeuPheArgAsnLeuAspValValProIleAsp
 AGAGATAAGATTGAGAAAGAAAATGCACTTTTTTCGTAACCTTGATGTAGTACCAATAGAT
 AsnAlaSerThrThrThrAsnTyrThrAsnTyrArgLeuIleHisCysAsnArgSerVal 203
 AATGCTAGTACTACTACCAACTATACCAACTATAGGTTGATACATTGTAACAGATCAGTC
 6388 IleThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCysThrPro
 ATTACACAGGCCTGTCCAAAGGTATCATTGTAGCCAATTCCCATACATTATTGTACCCCG
 AlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyLysGlyProCysThr 243 E
 GCTGGTTTTGCGATTCTAAAGTGTAATAATAAACGTTCAATGGAAAAGGACCATGTACA
 6508 AsnValSerThrValGlnCysThrHisGlyIleArgProIleValSerThrGlnLeuLeu
 AATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAATAGTGTCAACTCAACTGCTG
 LeuAsnGlySerLeuAlaGluGluGluValValIleArgSerAspAsnPheThrAsnAsn 283 N
 TTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTACGAACAAAT
 6628 AlaLysThrIleIleValGlnLeuAsnGluSerValAlaIleAsnCysThrArgProAsn
 GCTAAAACCATAATAGTACAGCTGAATGAATCTGTAGCAATTAACGTACAAGACCCAAAC
 AsnAsnThrArgLysSerIleTyrIleGlyProGlyArgAlaPheHisThrThrGlyArg 323 V
 AACCAATACAAGAAAAAGTATCTATATAGGACCAGGGAGAGCATTTCATACAACAGGAAGA
 6748 IleIleGlyAspIleArgLysAlaHisCysAsnIleSerArgAlaGlnTrpAsnAsnThr
 ATAATAGGAGATATAAGAAAAGCACATTGTAACATTAGTAGAGCACAATGGAATAACACT
 LeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsnAsnLysThrIleValPhe 363
 TTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGGAATAATAAAACAATAGTCTTT
 6868 AsnGlnSerSerGlyGlyAspProGluIleValMetHisSerPheAsnCysArgGlyGlu
 AATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTAGAGGGGAA
 PhePheTyrCysAsnThrThrGlnLeuPheAsnAsnThrTrpArgLeuAsnHisThrGlu 403
 TTTTCTACTGTAATACAACACAACACTGTTTAATAATACATGGAGGTTAAATCACACTGAA
 6988 GlyThrLysGlyAsnAspThrIleIleLeuProCysArgIleLysGlnIleIleAsnMet
 GGAACATAAGGAAATGACACAATCATACTCCCATGTAGAATAAAACAAATTATAAACATG
 TrpGlnGluValGlyLysAlaMetTyrAlaProProIleGlyGlyGlnIleSerCysSer 443
 TGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATTTGGAGGACAAATTAGTTGTTCA
 7108 SerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThrAsnValThrAsnAspThr
 TCAATATTACAGGGCTGCTATTAAACAAGAGATGGTGGTACAAATGTAACATAATGACACC
 GluValPheArgProGlyGlyGlyAspMetArgAspAsnTrpArgSerGluLeuTyrLys 483
 GAGGTCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA

FIG. 5F

TyrLysValIleLysIleGluProLeuGlyIleAlaProThrLysAlaLysArgArgVal
 7228 TATAAAGTAATAAAAAATTGAACCATTAGGAATAGCACCCACCAAGGCAAAGAGAAGAGTG
 ValGlnArgGluLysArgAlaValGlyIleValGlyAlaMetPheLeuGlyPheLeuGly 523
 GTGCAGAGAGAAAAAGAGCAGTGGGAATAGTAGGAGCTATGTTCTTGGGTTCTTGGGA
 AlaAlaGlySerThrMetGlyAlaValSerLeuThrLeuThrValGlnAlaArgGlnLeu
 7348 GCAGCAGGAAGCACTATGGGCGCAGTGTCTTACGCTGACGGTACAGGCCAGACAATTA
 LeuSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGlnHis 563
 TTGTCTGGTATAGTGCAACAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAACAT
 LeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaValGlu
 7468 CTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCTTGGCTGTGGAA
 ArgTyrLeuArgAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIleCys 603
 AGATACCTAAGGGATCAACAGCTCCTAGGGATTTGGGGTTGCTCTGGAAAACCTCATTTGC
 ThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluAspIleTrpAsp
 7588 ACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAAGACATTTGGGAT
 AsnMetThrTrpMetGlnTrpGluArgGluIleAspAsnTyrThrAsnThrIleTyrThr 643
 AACATGACCTGGATGCAGTGGGAAAGAGAAATTGACAATTACACAAACACAATATACACC
 LeuLeuGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeuAsp
 7708 TTACTIONGAGAATCGCAGAACCAACAAGAAAAGAATGAACAAGAATTATTAGAATTGGAT
 LysTrpAlaSerLeuTrpAsnTrpPheSerIleThrAsnTrpLeuTrpTyrIleLysIle 683
 AAGTGGGCAAGTTTGTGGAATTGGTTTAGCATAACAACTGGCTGTGGTATATAAAGATA
 PheIleMetIleValGlyGlyLeuValGlyLeuArgIleValPheAlaValLeuSerIle E
 7828 TTCATAATGATAGTAGGAGGCTTGGTAGGTTAAGAATAGTTTTTGTCTGTGCTTTCTATA
 ValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrArgLeuProValPro 723
 GTGAATAGAGTTAGGCAGGGATACTCACCATTGTCTTTTACAGACCCGCCTCCAGTCCCG N
 ArgGlyProAspArgProAspGlyIleGluGluGluGlyGlyGluArgAspArgAspArg
 7948 AGGGGACCCGACAGGCCCGACGGAATCGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGA
 SerValArgLeuValAspGlyPheLeuAlaLeuIleTrpGluAspLeuArgSerLeuCys 763 V
 TCCGTTGATTAGTGGATGGATTCTTAGCACTTATCTGGGAAGATCTGCGGAGCCTGTGC
 LeuPheSerTyrArgArgLeuArgAspLeuLeuLeuIleAlaAlaArgThrValGluIle
 8068 CTCTTCAGCTACCGCCGCTTGAGAGACTTACTCTTGATTGCAGCGAGGACTGTGGAAATT
 LeuGlyHisArgGlyTrpGluAlaLeuLysTyrTrpTrpSerLeuLeuGlnTyrTrpIle 803
 CTGGGGCACAGGGGGTGGGAAGCCCTCAAATATTGGTGGAGTCTCCTGCAGTATTGGATT
 GlnGluLeuLysAsnSerAlaValSerTrpLeuAsnAlaThrAlaIleAlaValThrGlu
 8188 CAGGAACATAAGAATAGTGCTGTTAGCTGGCTCAACGCCACAGCTATAGCAGTAAGTGG
 GlyThrAspArgValIleGluValAlaGlnArgAlaTyrArgAlaIleLeuHisIleHis 843
 GGGACAGATAGGGTTATAGAAGTAGCACAAAGAGCTTATAGAGCTATTCTCCACATACAT
 ArgArgIleArgGlnGlyLeuGluArgLeuLeuLeuOC
 8308 AGAAGAATTAGACAGGGCTTGGAAAGGCTTTTGTCTATAAGATGGGTGGCAAGTGGTCAAA
 ACGTAGTATGGGTGGATGGTCTGCTATAAGGGAAAGAATGAGACGAGCTGAGCCACGAGC

FIG. 5G

8428 TGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTGGAAAAACATGGAGCAAT
 CACAAGTAGCAATACAGCAGCTACTAATGCTGATTGTGCCTGGCTAGAAGCACAAGAGGA
 8548 GGAAGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGACTTACAAGGC
 AGCTTTAGATATTAGCCACTTTTTAAAAGAAAAGGGGGGA ^{U3 →} CTGGAAGGGCTAATTTGGT
 8667 CCCAAAGAAGACAAGAGATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTG
 ATTGGCAGAATTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCT
 8787 TCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCAATGAAGGAGAGAACAACA
 GCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACGCGGAGAAAGAAGTGTTAGTGT
 8907 GGAGGTTTGACAGCAAAGTAGCATTTTCATCACATGGCCCCGAGAGCTGCATCCGGAGTACT
 ACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGA
 9027 GGCCTGGCCTGGGCGGGACTGGGGAGTGGCGTCCCTCAGATGCTGCATATAAGCAGCTGC
 TTTTTGCCTGTACTG ^{← U3 R →} GGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC
 9146 TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCA ^{← R U5 →} AGTAGT
 GTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGT
 9265 GTGGAAAAATCTCTAGCAG ^{← U5}

L
T
R

FIG. 5H

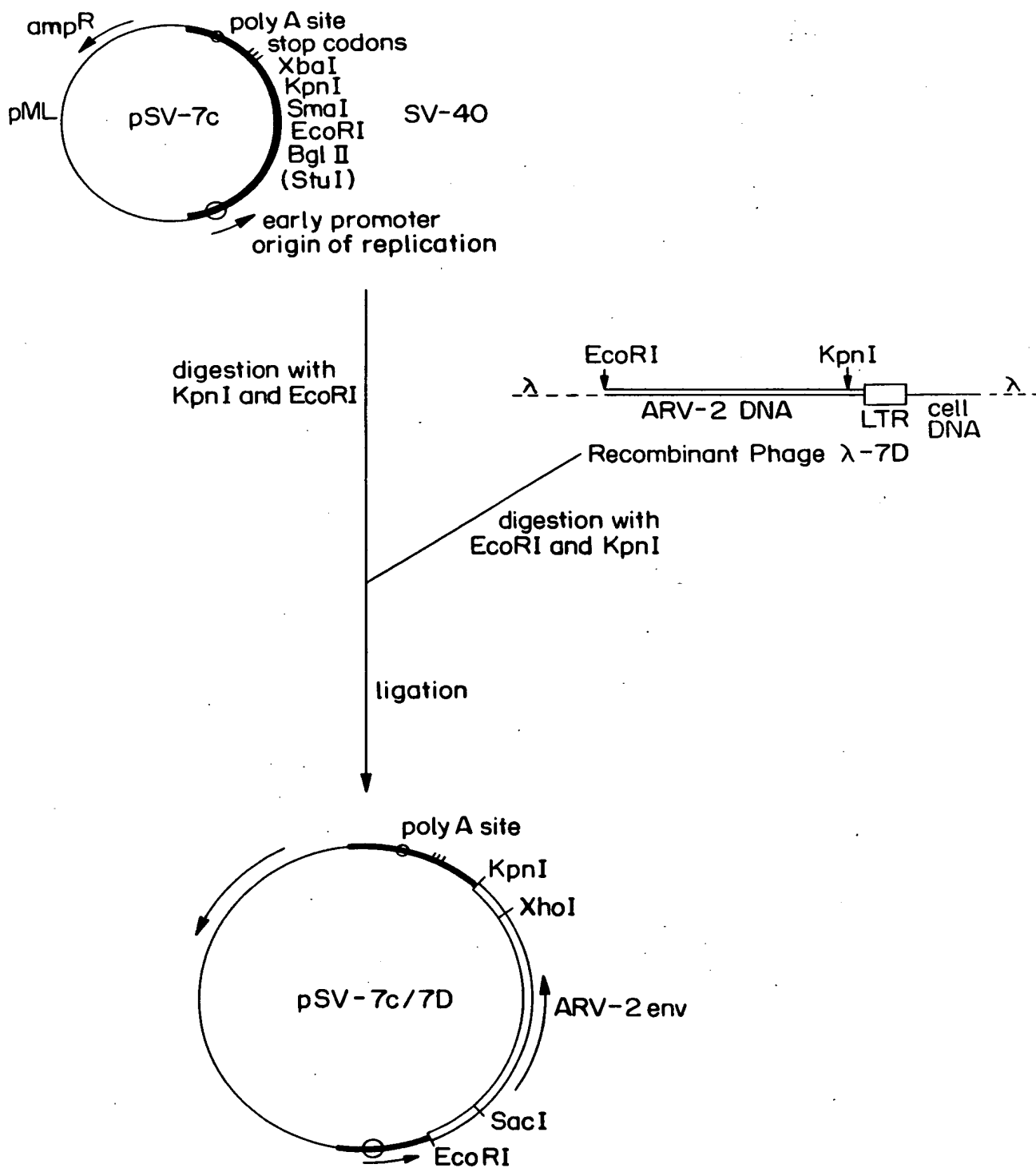
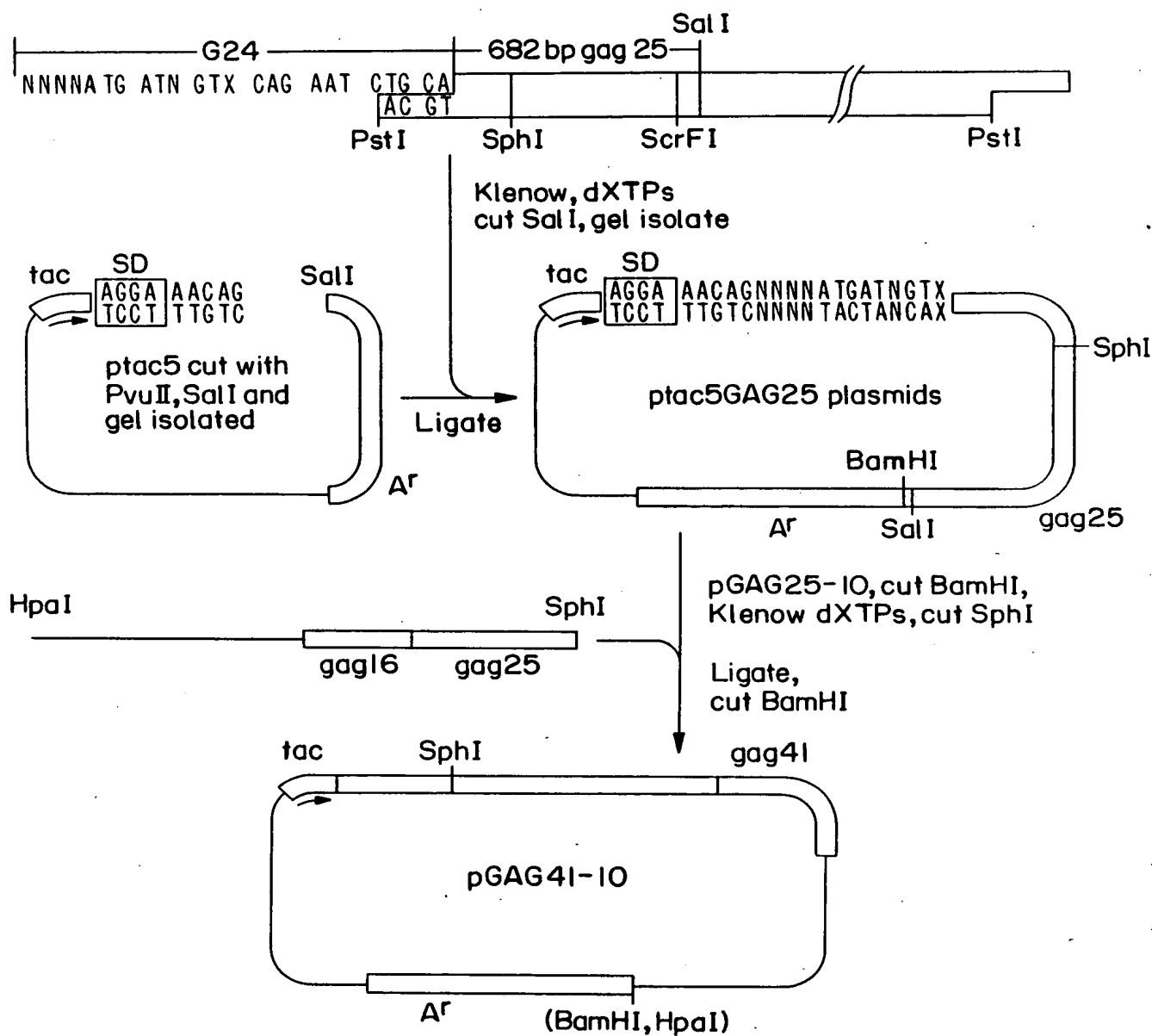


FIG. 6

FIG. 7





*

ptac 5 Promotor	Met Ile Val ATGATCGTA
-----------------	--------------------------

748 GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrp
CAGAATCTGCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGG

VallysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu 181
GTAAAAGTAGTAGAAGAAAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTTTCAGCATTA

868 SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGln
TCAGAAGGAGCCACCCACAAGATTAAACACCATGCTAAACACAGTGGGGGGACATCAA

AlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal 221
GCAGCCATGCAAATGTTAAAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGAGTG

988 HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAsp
CATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAC

IleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro 261
ATAGCAGGAACACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT

1108 IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArg
ATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGA

MetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp 301
ATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCCTTAGAGAT

1228 TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsn
TATGTAGACCGGTTCTATAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAAT

TrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys 341
TGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAGATTGTAAGACTATTTTAAAA

1348 AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGly
GCATTGGGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGA

ProGlyHisLysAlaArgValLeu Stop Stop
CCCGGGCATAAAGCAAGAGTTTTGTGATAG

ptac 5

FIG. 8

ptac 5 Promotor

MetIleVal 141
ATGATCGTA

748 GlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrp
 CAGAATCTGCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAAGCTTTAAATGCATGG
 ValLysValValGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeu 181
 GTAAAGTAGTAGAAGAAAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTTTCAGCATTA G
 868 SerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGln
 TCAGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAA
 AlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgVal 221
 GCAGCCATGCAAATGTTAAAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGAGTG
 988 HisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAsp A
 CATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAC
 IleAlaGlyThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPro 261
 ATAGCAGGAAGTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCT
 1108 IleProValGlyGluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArg G
 ATCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGA
 MetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAsp 301
 ATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGAT
 1228 TyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAlaSerGlnAspValLysAsn
 TATGTAGACCGGTTCTATAAACTCTAAGAGCCGAACAAGCTTCACAGGATGTAAAAAAT
 TrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLys 341
 TGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAGATTGTAAGACTATTTTAAAA
 1348 AlaLeuGlyProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGly
 GCATTGGGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGA
 ProGlyHisLysAlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaAsn 381
 CCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCCATGAGCCAAGTAACAAATCCAGCTAAC
 1468 IleMetMetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCys
 ATAATGATGCAGAGAGGCAATTTTAGGAACCAAGAAAGACTGTTAAGTGTTTCAATTGT
 GlyLysGluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpArg 421
 GGCAAAGAAGGGCACATAGCCAAAAATTGCAGGGCCCTAGGAAAAAGGGCTGTTGGAGA
 CysGlyArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGly
 1588 TGTGGAAGGGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGG PhePheArgG
 LysIleTrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluPro 461
 luAspLeuAlaPheLeuGlnGlyLysAlaArgGluPheSerSerGluGlnThrArgAla 23
 AAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCA
 ThrAlaProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLys
 1708 AsnSerProThrArgArgGluLeuGlnValTrpGlyGlyGluAsnAsnSerLeuSerGluA
 ACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACCTCCCTCTCAGAAG
 GlnGluProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsn 501
 laGlyAlaAspArgGlnGlyThrValSerPheAsnPheProGlnIleThrLeuTrpGln 63
 CAGGAGCCGATAGACAAGGAAGTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAAC

FIG. 9A

AspProSerSerGlnOC
 ArgProLeuValThrIleArgIleGlyGlyGlnLeuLysGluAlaLeuLeuAspThrGlyA
 1828 GACCCCTCGTCACAATAAGGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAG

 laAspAspThrValLeuGluGluMetAsnLeuProGlyLysTrpLysProLysMetIle 103
 CAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAG

 GlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGlnIleProValGluIleCysG
 1948 GGGGAATTGGAGGTTTTATCAAAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTG

 lyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArg 143
 GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA

 AsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValP
 2068 ATCTGTTGACTCAGATTGGTTGTACTTTAAATTTCCCATTAGTCCTATTGAACTGTAC

 roValLysLeuLysProGlyMetAspGlyProLysValLysGlnTrpProLeuThrGlu 183
 CAGTAAATTAAGCCAGGAATGGATGGCCCAAAAGTTAAGCAATGGCCATTGACAGAAG

 GluLysIleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIleSerL
 2188 AAAAAATAAAGCATTAGTAGAGATATGTACAGAAATGGAAAAGGAAGGGAAAAATTTCAA

 ysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSer 223
 AAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCTATAAAGAAAAAAGACAGTA

 ThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrpG
 2308 CTAATGGAGAAACTAGTAGATTTTCAGAGAAGCTTAATAAAGAACTCAAGACTTCTGGG

 luValGlnLeuGlyIleProHisProAlaGlyLeuLysLysLysLysSerValThrVal 263
 AAGTTCAGTTAGGAATACCAACCCCGCAGGGTTAAAAAAGAAAAATCAGTAACAGTAT

 LeuAspValGlyAspAlaTyrPheSerValProLeuAspLysAspPheArgLysTyrThrA
 2428 TGGATGTGGGTGATGCATACTTTTCAGTTCCCTTAGATAAAGACTTTAGAAAGTATACTG

 laPheThrIleProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnVal 303
 CATTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGC

 LeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuG
 2548 TGCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAG

 luProPheArgLysGlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyr 343
 AGCCTTTTAGAAAACAGAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATG

 ValGlySerAspLeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHisL
 2668 TAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAACTGAGACAGCATC

 euLeuArgTrpGlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeu 383
 GTTTGAGGTGGGGATTTACCACACCAGACAAAAACATCAGAAAGAACCTCCATTCTTT

 TrpMetGlyTyrGluLeuHisProAspLysTrpThrValGlnProIleMetLeuProGluL
 2788 GGATGGGTTATGAATCCATCCTGATAAATGGACAGTACAGCCTATAATGCTGCCAGAAA

 ysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSer 423
 AAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAAATTGAATTGGGCAAGTC

 GlnIleTyrAlaGlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyThrLysAlaL
 2908 AGATTTTATGCAGGGATTAAAGTAAAGCAGTTATGTAAACTCCTTAGAGGAACCAAAGCAC

 euThrGluValIleProLeuThrGluGluAlaGluLeuGluLeuAlaGluAsnArgGlu 463
 TAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAGA

FIG. 9B

IleLeuLysGluProValHisGluValTyrTyrAspProSerLysAspLeuValAlaGluI
 3028 TTCTAAAAGAACCAGTACATGAAGTATATTATGACCCATCAAAAGACTTAGTAGCAGAAA
 leGlnLysGlnGlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluProPheLysAsn 503
 TACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTAAAAATC
 LeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuT
 3148 TGAAAACAGGAAAGTATGCAAGGATGAGGGGTGCCACACTAATGATGTAAAACAGTT
 hrGluAlaValGluLysValSerThrGluSerIleValIleTrpGlyLysIleProLys 543
 ptac 5

FIG. 9C

ARV GAG p16 - synthetic Parts A and B

5' arv 234 3' MetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLys
TATTATGCAAAGAGGTAACCTTCAGGAATCAAAGAAAGACCGTTAAGTGTTCACCTGTGGTAAG
ATAATACGTTTCTCCATTGAAGTCCTTAGTTTCTTTCTGGCAATTCACAAAGTTGACACCATTC
3' arv 235 5'
10 mnl1, 23 hinf1,
63 GluGlyHisIleAlaLysAsnCysArgAlaProArgLysLysAlaCysTrpArgCysGly
GAAGGTCACATCGCTAAGAACTGTAGAGCTCCAAGAAAGAAGGCTTGTTGGAGATGTGGT
CTTCCAGTGTAGCGATTCTTGACATCTCGAGGTTCTTTCTTCCGAACAACCTCTACACCA
76 dde1, 88 ban2 hgiA hgiJ11 sac1 sduI, 89 alu1,
123 ArgGluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle
AGAGAAGGTCACCAAATGAAGGACTGTACCGAAAGACAAGCTAACTTCTTGGGTAAAGATC
TCTCTTCCAGTGGTTTACTTCTGACATGGCTTTCTGTTCGATTGAAGAACCCATTCTAG
129 bstE2, 131 hph, 148 rsaI, 161 alu1, 178 bgl11 xho2, 179
sau3a,
183 TrpProSerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAla
TGGCCATCTTACAAGGGTAGACCAAGGTAACCTTCTTGCAATCCAGACCAGAACCAACCGCT
ACCGGTAGAATGTTCCCATCTGGTCCATTGAAGAACGTTAGGTCTGGTCTTGTTGGTGGCGA
183 ball cfr1 haeI, 184 hae111, 199 acc1, 204 apy1 ecor11 sc
rF1,
243 ProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGlnLysGlnGlu
CCACCTGAAGAAAGTTTCAGGTTTCGGTGAAGAAAAGACCACCCCATCTCAAAAGCAAGAA
GGTGGACTTCTTTCAAAGTCCAAGCCACTTCTTTCTGGTGGGGTAGAGTTTTCGTTCTT
249 mbo11, 267 hph, 270 mbo11,
303 ProIleAspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsnAspPro
CCAATCGACAAGGAATTGTACCCATTGACCTCTTTGAGATCCTTGTTTCGGTAACGATCCC
GGTTAGCTGTTCTTAACATGGGTAACCTGGAGAACTCTAGGAACAAGCCATTGCTAGGG
307 taq1, 320 rsaI, 331 mnl1, 339 xho2, 340 sau3a, 357 sau3a,
361 mnl1, 362 aval xho1,
363 SerSerGlnOP AM
TCGAGCCAATGATAG
AGCTCGGTTACTATCAGCT
363 taq1, 377 acc1 hind11 sal1

FIG. 10

PYK Promoter

 MetSer
 ATGTCT

ArgIleAspCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProVal 51
 AGAATCGAT GTAGTGCTACAGAAAAATTGTGGGTCACAGTTTATTATGGAGTACCTGTG

5908 TrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAlaArgAlaTyrAspThrGlu
 TGGAAAGAAGCAACTACCACTCTATTTTGTGCATCAGATGCTAGAGCATATGATACAGAG

ValHisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGluVal 91
 GTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTA

6028 ValLeuGlyAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGlnMet
 GTATTGGGAAATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAGATG

GlnGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysValLysLeuThrPro 131
 CAGGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCA

6148 LeuCysValThrLeuAsnCysThrAspLeuGlyLysAlaThrAsnThrAsnSerSerAsn
 CTCTGTGTTACTTTAAATTGCACTGATTTGGGGAAGGCTACTAATACCAATAGTAGTAAT

TrpLysGluGluIleLysGlyGluIleLysAsnCysSerPheAsnIleThrThrSerIle 171
 TGGAAAGAAGAAATAAAAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGCATA

6268 ArgAspLysIleGlnLysGluAsnAlaLeuPheArgAsnLeuAspValValProIleAsp
 AGAGATAAGATTTCAGAAAGAAAATGCACCTTTTCGTAACTTGATGTAGTACCAATAGAT

AsnAlaSerThrThrThrAsnTyrThrAsnTyrArgLeuIleHisCysAsnArgSerVal 211
 AATGCTAGTACTACTACCAACTATACCAACTATAGGTTGATACATTGTAACAGATCAGTC

6388 IleThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCysThrPro
 ATTACACAGGCCTGTCCAAAGGTATCATTGAGCCAATTCCCATACATTATTGTACCCCG

AlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyLysGlyProCysThr 251
 GCTGGTTTTGCGATTCTAAAGTGAATAATAAAACGTTCAATGGAAAAGGACCATGTACA

FIG. 11A

6508 AsnValSerThrValGlnCysThrHisGlyIleArgProIleValSerThrGlnLeuLeu
 AATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAATAGTGTCAACTCAACTGCTG
 LeuAsnGlySerLeuAlaGluGluGluValValIleArgSerAspAsnPheThrAsnAsn 291
 TTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTCACGAACAAT
 6628 AlaLysThrIleIleValGlnLeuAsnGluSerValAlaIleAsnCysThrArgProAsn
 GCTAAAACCATAATAGTACAGCTGAATGAATCTGTAGCAATTAAGTGTACAAGACCCAAC
 AsnAsnThrArgLysSerIleTyrIleGlyProGlyArgAlaPheHisThrThrGlyArg 331
 AACAAACAAGAAAAAGTATCTATATAGGACCAGGGAGAGCATTTTCATACAACAGGAAGA
 6748 IleIleGlyAspIleArgLysAlaHisCysAsnIleSerArgAlaGlnTrpAsnAsnThr
 ATAATAGGAGATATAAGAAAAAGCACATTGTAACATTAGTAGAGCACAATGGAATAACACT
 LeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsnAsnLysThrIleValPhe 371
 TTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGGAATAATAAAACAATAGTCTTT
 6868 AsnGlnSerSerGlyGlyAspProGluIleValMetHisSerPheAsnCysArgGlyGlu
 AATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTAGAGGGGAA
 PhePheTyrCysAsnThrThrGlnLeuPheAsnAsnThrTrpArgLeuAsnHisThrGlu 411 E
 TTTTCTACTGTAATACAACACAACACTGTTTAATAATACATGGAGGTTAAATCACACTGAA
 6988 GlyThrLysGlyAsnAspThrIleIleLeuProCysArgIleLysGlnIleIleAsnMet
 GGAAGTAAAGGAAATGACACAATCATACTCCCATGTAGAATAAAACAAATTATAAACATG
 TrpGlnGluValGlyLysAlaMetTyrAlaProProIleGlyGlyGlnIleSerCysSer 451 N
 TGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATTTGGAGGACAAATTAGTTGTTCA
 7108 SerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThrAsnValThrAsnAspThr
 TCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTACAAATGTAACATAATGACACC
 GluValPheArgProGlyGlyGlyAspMetArgAspAsnTrpArgSerGluLeuTyrLys 491 V
 GAGGTCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
 7228 TyrLysValIleLysIleGluProAsnSerValSer
 TATAAAGTAATAAAAAATTGAACCAATTCGGTATCTTGA

PYK Terminator

FIG. 11B

Nucleotide positions relative to Figure 5.	1	MetIleAspLysAlaGlnGluGluHisGluLysTyrHisSerAsnTrp AGGXAACAG:::ATGAT:GA:AAGGCACAAGAAGAACATGAGAAATATCACAGTAATTGG TCCXTTGTCT:::TACTA:CT:TTCCGTGTTCTTCTTGTACTCTTTATAGTGTCTTAATACC
		32 mbo11, 38 nla111,
3820	62	ArgAlaMetAlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSer AGAGCCATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGC TCTCGGTACCGATCACTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTCTG
		66 nco1, 67 nla111, 118 nspBII pvu11, 119 alu1,
3880	122	CysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGly TGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGA ACACTATTTACAGTCTGATTTTCTCTTCGGTACGTACCTGTTCTATCTGACATCAGGTCTCT
		135 alu1, 151 nla111, 152 ns11 ava3, 155 nla111, 164 acc1, 1 76 apy1 bstXI ecor11 scrF1,
3940	182	IleTrpGlnLeuAspCysThrHisLeuGluGlyLysIleIleLeuValAlaValHisVal ATATGGCAACTAGATTGTACACATCTAGAAGGAAAAATTATCCTGGTAGCAGTTCATGTA TATACCGTTGATCTAACTGTGTAGATCTTCTTTTTTAATAGGACCATCGTCAAGTACAT
		198 rsa1, 205 xba1, 223 apy1 ecor11 scrF1, 236 nla111,
4000	242	AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyr GCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAGACAGGGCAGGAAACAGCATAT CGGTCACCTATATATCTTCGTCTTCAATAAGGTCGTCTCTGTCCCGTCTTTGTCGTATA
		263 xmn1,
4060	302	PheLeuLeuLysLeuAlaGlyArgTrpProValLysThrIleHisThrAspAsnGlySer TTTCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGC AAAGAGAATTTTAAATCGTCTTCTACCGGTCATTTTTGTTATGTATGTCTGTTACCGTCTG
		321 mbo11, 326 bal1 cfr1 hae1, 327 hae111, 357 bbv fnu4h1,
4120	362	AsnPheThrSerThrThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPhe AATTTCCACGACTACTACGGTTAAGGCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTT TTAAAGTGGTCTATGATGCCAATTCCGGCGGACAACCAACCCGTCCTTAGTTCGTCTTAAA
		366 hph, 371 sca1, 372 rsa1, 385 hae111, 386 fnu4h1 nsb11, 4 05 bin1, 406 dpn1 sau3a,
4180	422	GlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMetAsnAsnGluLeuLys GGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAAG CCGTAAGGGATGTTAGGGGTTTCAGTTCTCATCATCTTAGATACTTATTACTTAATTTCT
		423 bsm1, 458 hinf1,
4240	482	LysIleIleGlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAla AAAATTATAGGACAGGTAAGAGATCAGGCTGAACACCTTAAGACAGCAGTACAAATGGCA TTTTAATATCCTGTCCATTCTCTAGTCCGACTTGTGGAATTCTGTCTCATGTTTACCGT
		503 dpn1 sau3a, 518 afl111, 530 rsa1,
4300	542	ValPheIleHisAsnPheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArg GTATTCATCCACAATTTTAAAGAAAAGGGGGGATTGGGGGATACAGTGCAGGGGAAAGA CATAAGTAGGTGTTAAATTTTCTTTTCCCCCTAACCCCTATGTCACGTCCCCTTTCT
		547 fok1, 557 aha111,

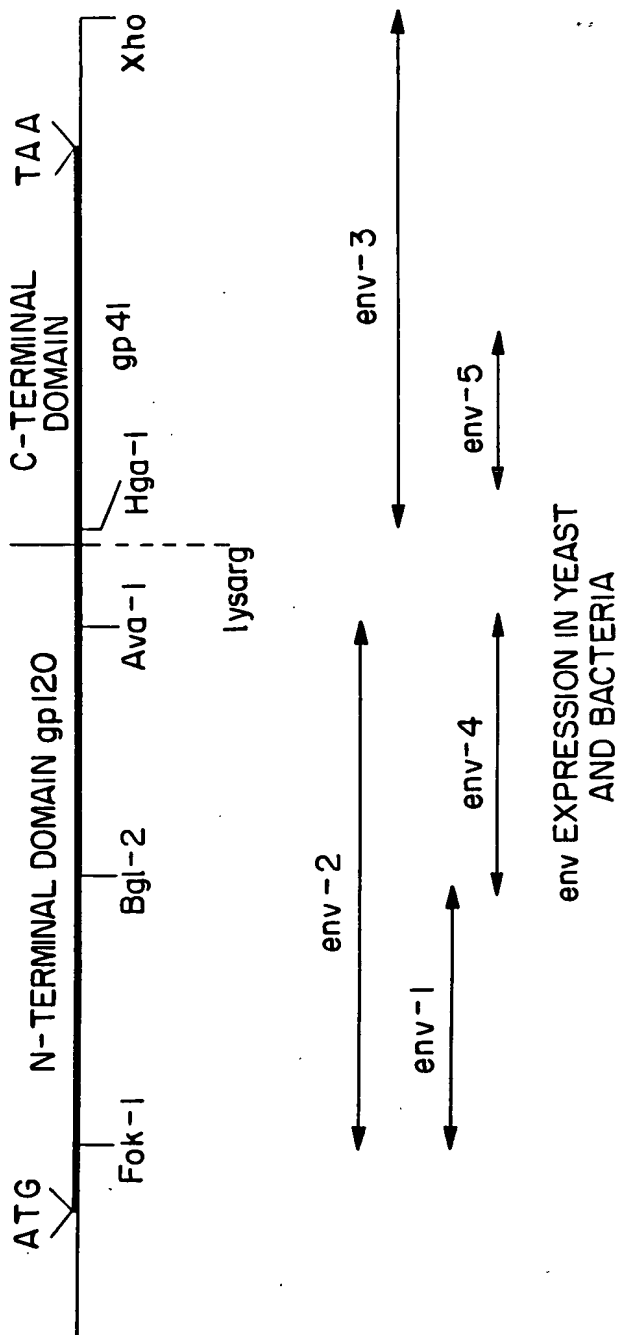
FIG. 12A

4360 602 IleValAspIleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLys
 ATAGTAGACATAATAGCAACAGACATACAACTAAAGAAGCAAAAGCAAAATTACAAAA
 TATCATCTGTATTATCGTTGTCTGTATGTTTGATTTCCTTGATGTTTTCGTTTAATGTTTT
 605 acc1,
 4420 662 IleGlnAsnPheArgValTyrTyrArgAspAsnLysAspProLeuTrpLysGlyProAla
 ATTCAAAATTTTCGGGTTTATTACAGGGACAACAAAGATCCCCCTTTGGAAAGGACCAGCA
 TAAGTTTTAAAAGCCCCAAATAATGTCCTGTTGTTTCTAGGGGAAACCTTTCCTGGTCGT
 697 xho2, 698 dpn1 sau3a, 713 asu1 ava2,
 4480 722 LysLeuLeuTrpLysGlyGluGlyAlaValValIleGlnAspAsnSerAspIleLysVal
 AAGCTTCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTA
 TTCGAAGAGACCTTTCCTTCCCCGTCATCATTATGTTCTATTATCACTGTATTTTCAT
 722 hind111, 723 alu1, 737 hph,
 4540 782 ValProArgArgLysAlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAsp
 GTGCCAAGAAGAAAAGCAAAATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGAT
 CACGGTTCTTCTTTTCGTTTTTAGTAATCCCTAATACCTTTTGTCTACCGTCCACTACTA
 789 mbo11, 833 hph,
 4600 842 CysValAlaSerArgGlnAspGluAspAM
 TGTGTGGCAAGTAGACAGGATGAGGATTAGTCGACGGAATTCTTTAGTAAACACC
 ACACACCGTTTCATCTGTCCTACTCCTAATCAGCTGCCCTTAAGAAATCATTTTGTGG
 852 acc1, 859 fok1, 863 mnl1, 871 acc1 hind11 sal1, 872 taq1
 , 878 ecor1,

FIG. 12B

FIG.13

HYDROPHOBIC REGIONS



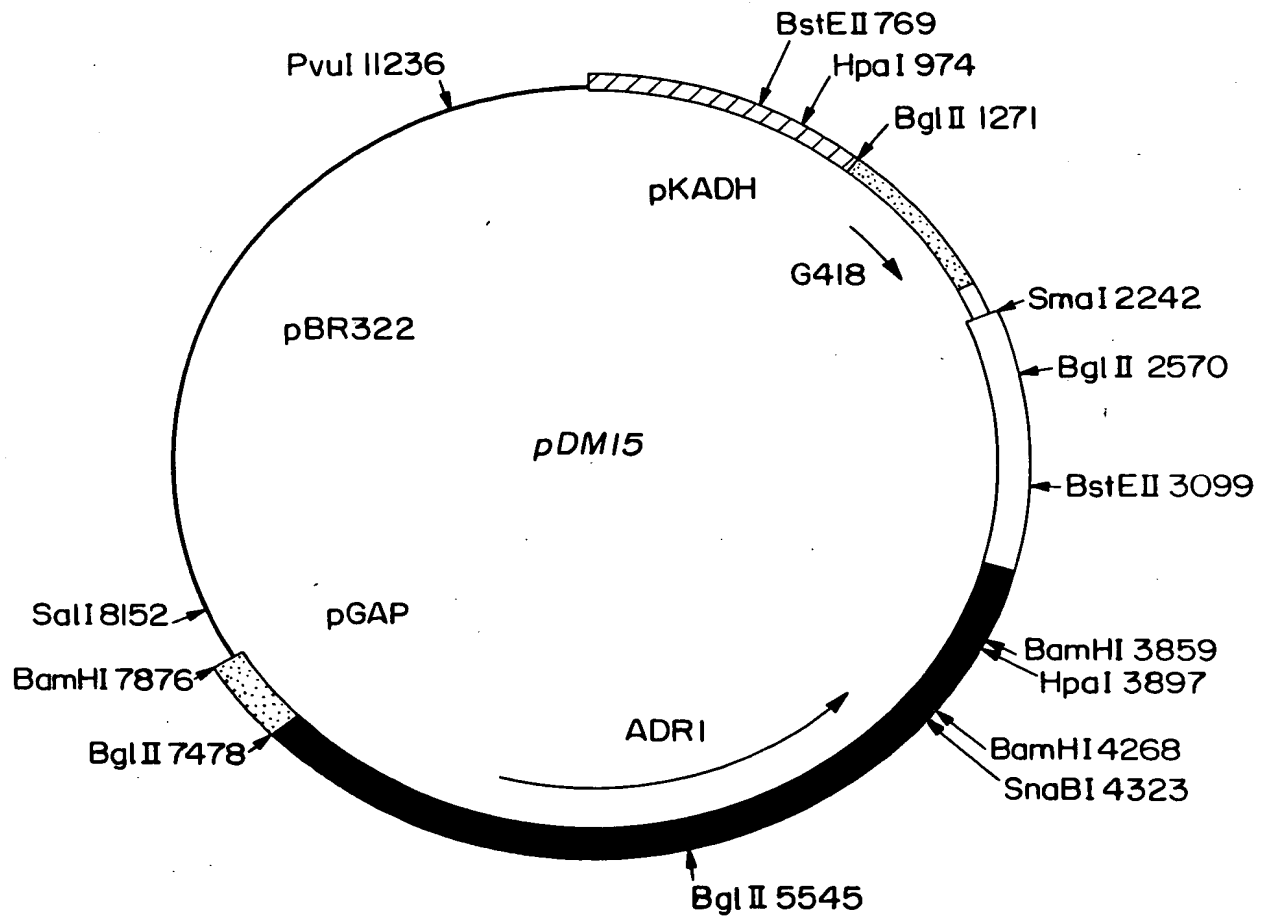


FIG.14

SOD

1 MetAlaThrLysAlaValCysValLeuLysGlyAspGlyProValGlnGlyIleIleAsn
 CATGGCGACGAAGGCCGTGTGCGTGCTGAAGGGCGACGGCCAGTGCAGGGCATCATCAAT
 CGCTGCTTCCGGCACACGCACGACTTCCCGCTGCCGGGTACAGTCCCGTAGTAGTGA

62 PheGluGlnLysGluSerAsnGlyProValLysValTrpGlySerIleLysGlyLeuThr
 TTCGAGCAGAAGGAAAGTAATGGACCAGTGAAGGTGTGGGGAAGCATTAAAGGACTGACT
 AAGCTCGTCTTCTTTTATTACCTGGTCACTTCCACACCCCTTCGTAATTTCTGACTGA

122 GluGlyLeuHisGlyPheHisValHisGluPheGlyAspAsnThrAlaGlyCysThrSer
 GAAGGCCTGCATGGATTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCACT
 CTTCCGGACGTACCTAAGGTACAAGTACTCAAACCTCTATTATGTCTCCGACATGGTCA

182 AlaGlyProHisPheAsnProLeuSerArgLysHisGlyGlyProLysAspGluGluArg
 GCAGGTCCTCACTTTAATCCTCTATCCAGAAAACACGGTGGGCCAAAGGATGAAGAGAGG
 CGTCCAGGAGTGAAATTAGGAGATAGGTCTTTTGTGCCACCCGGTTTCTACTTCTCTCC

242 HisValGlyAspLeuGlyAsnValThrAlaAspLysAspGlyValAlaAspValSerIle
 CATGTTGGAGACTTGGGCAATGTGACTGCTGACAAAGATGGTGTGGCCGATGTGTCTATT
 GTACAACCTCTGAACCCGTTACTGACGACTGTTTCTACCACACCCGGCTACACAGATAA

302 GluAspSerValIleSerLeuSerGlyAspHisCysIleIleGlyArgThrLeuValVal
 GAAGATTCTGTGATCTCACTCTCAGGAGACCATTGCATCATTGGCCGCACACTGGTGGTC
 CTTCTAAGACACTAGAGTGAGAGTCTCTGGTAACGTAGTAACCGGCGTGTGACCACCAG

362 HisGluLysAlaAspAspLeuGlyLysGlyGlyAsnGluGluSerThrLysThrGlyAsn
 CATGAAAAAGCAGATGACTTGGGCAAAGGTGGAAATGAAGAAAGTACAAAGACAGGAAAC
 GTACTTTTTCTGCTACTGAACCCGTTTCCACCTTACTTCTTTTCTGTTTCTGCTCTTTG

ENV 5B

422 AlaGlySerArgLeuAlaCysGlyValIleGlyIleAlaMetAlaIleGluAlaGlnGln
 GCTGGAAGTCGTTTGGCTTGTGGTGTAAATGGGATCGCCATGGCTATCGAAGCTCAACAA
 CGACCTTCAGCAAACCGAACACCACATTAACCCTAGCGGTACCGATAGCTTCGAGTTGTT

482 HisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaVal
 CACTTGCTGCAGTTGACCGTTTGGGGTATCAAGCAGTTGCAGGCTAGAGTTTTGGCTGTT
 GTGAACGACGTCAACTGGCAAACCCCATAGTTCGTCAACGTCCGATCTCAAACCGACAA

542 GluArgTyrLeuArgAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIle
 GAAAGATACTTGAGAGATCAACAATTGTTGGGTATCTGGGGTGTCTGTTAAGTTGATT
 CTTTCTATGAACCTCTAGTTGTTAACAACCCATAGACCCCAACAAGACCATTAACATA

602 CysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluAspIleTrp
 TGTACCACCGCTGTTCCCTGGAACGCTTCTTGGTCTAACAAGTCTTTGGAAGACATCTGG
 ACATGGTGGCGACAAGGGACCTTGCAGAAGAACAGATTGTTTCAGAAACCTTCTGTAGACC

662 AspAsnMetThrTrpMetGlnTrpGluArgGluIleAspAsnTyrThrAsnThrIleTyr
 GACAACATGACCTGGATGCAATGGGAAAGAGAAATCGACAACCTACACCAACACCATCTAC
 CTGTTGTACTGGACCTACGTTACCCCTTCTCTTTAGCTGTTGATGTGGTTGTGGTAGATG

722 ThrLeuLeuGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu
 ACCTTGTGGAGGAATCTCAAAACCAACAAGAAAGAACGAACAAGAATTGTTGGAATTG
 TGAACAACCTCCTTAGAGTTTGGTTGTTCTTTCTTGTCTTAAACAACCTTAAC

782 AspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrAsnTrpAM
 GACAAGTGGGCAAGCTTGTGGAACCTGGTTCTCTATCACCAACTGGTAG
 CTGTTACCCGTTTCGAACACCTTGACCAAGAGATAGTGGTTGACCATCAGCT

Translated Mol. Weight = 30414.22

FIG. 15

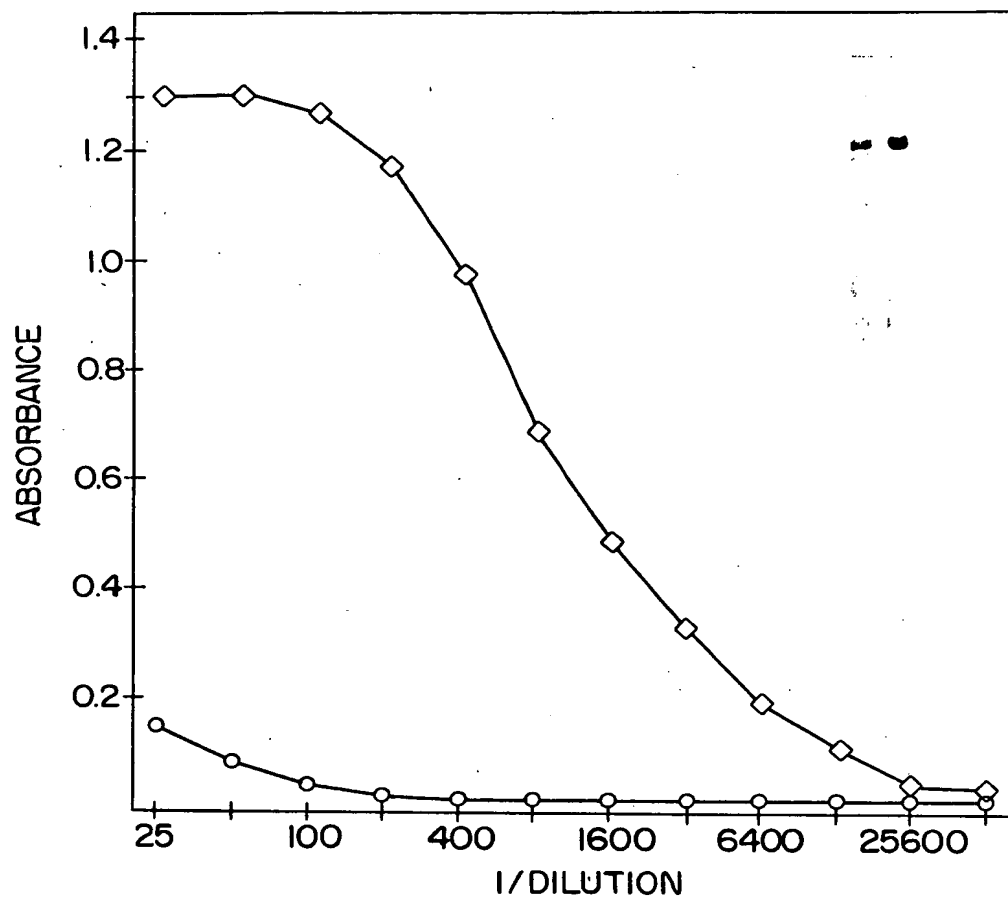


FIG. 16A

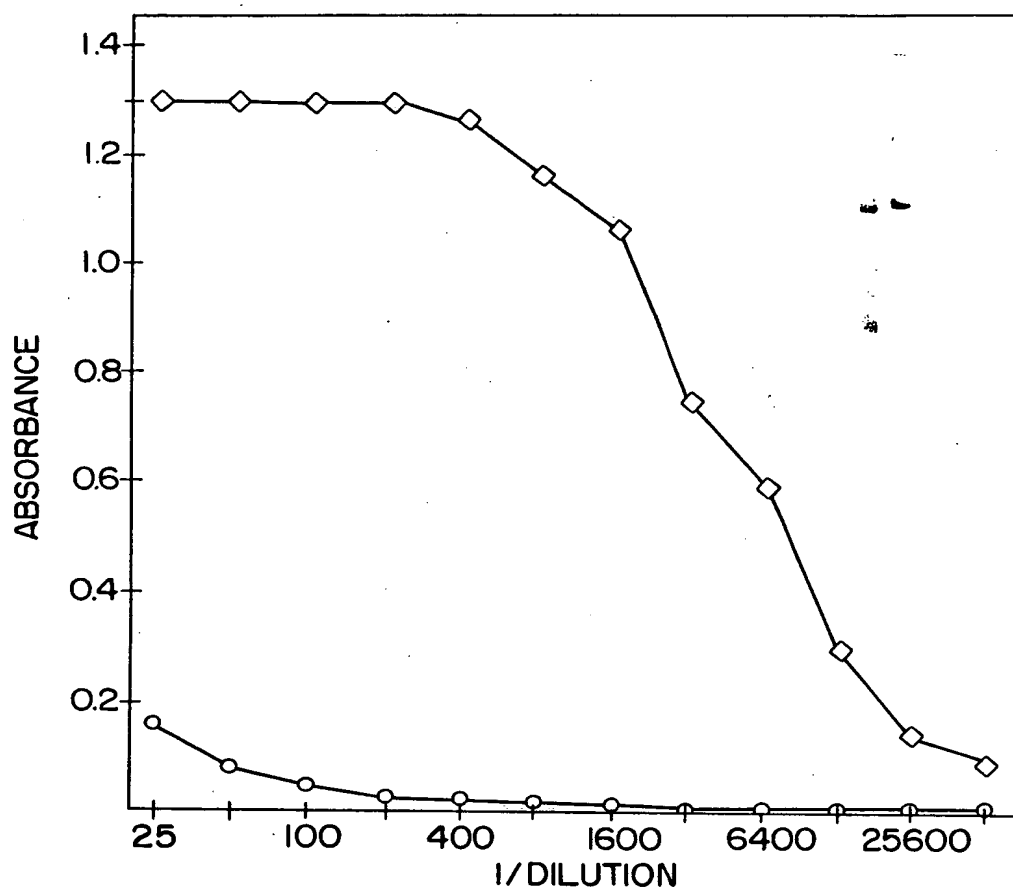


FIG. 16B

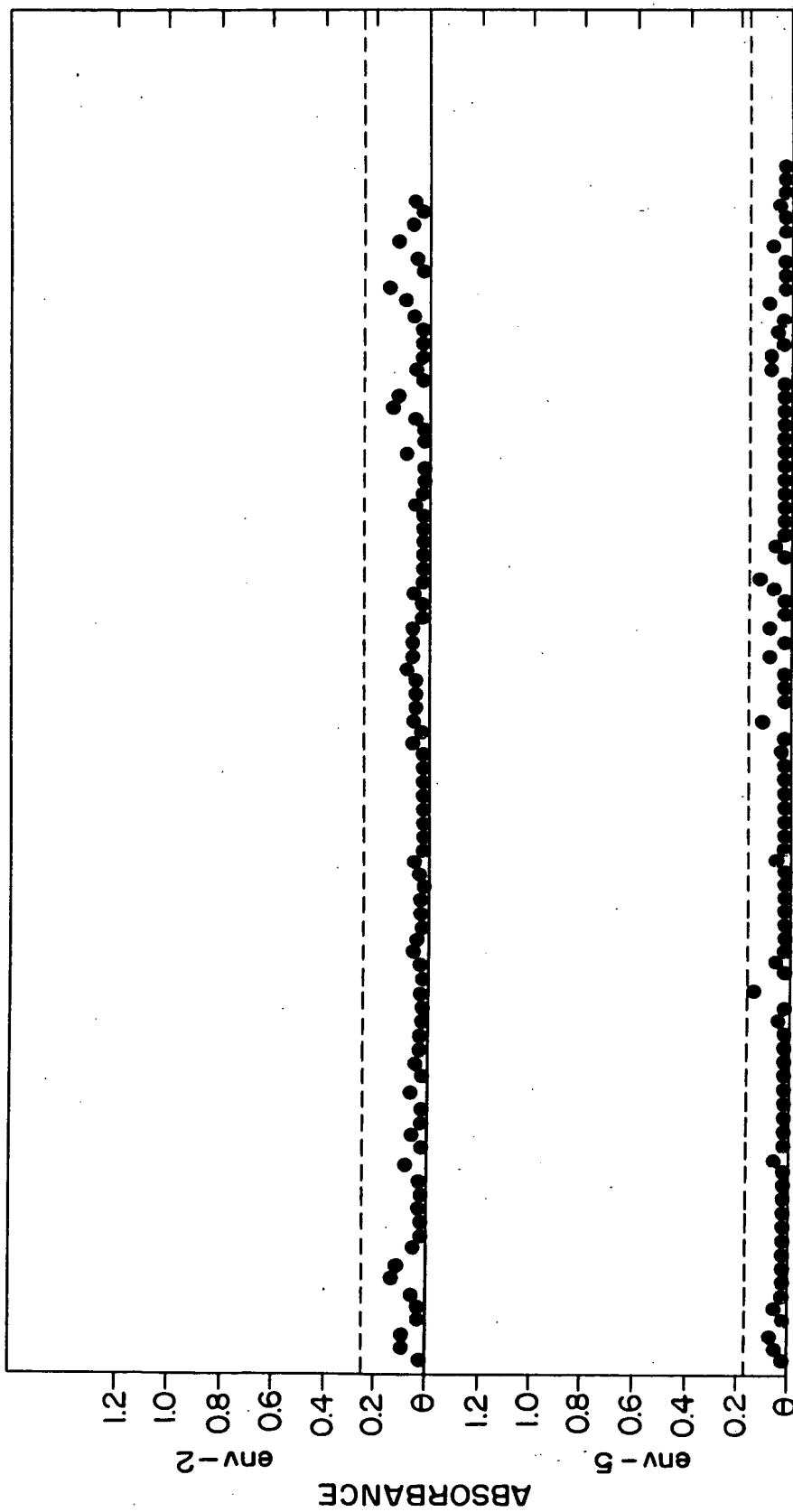


FIG.17

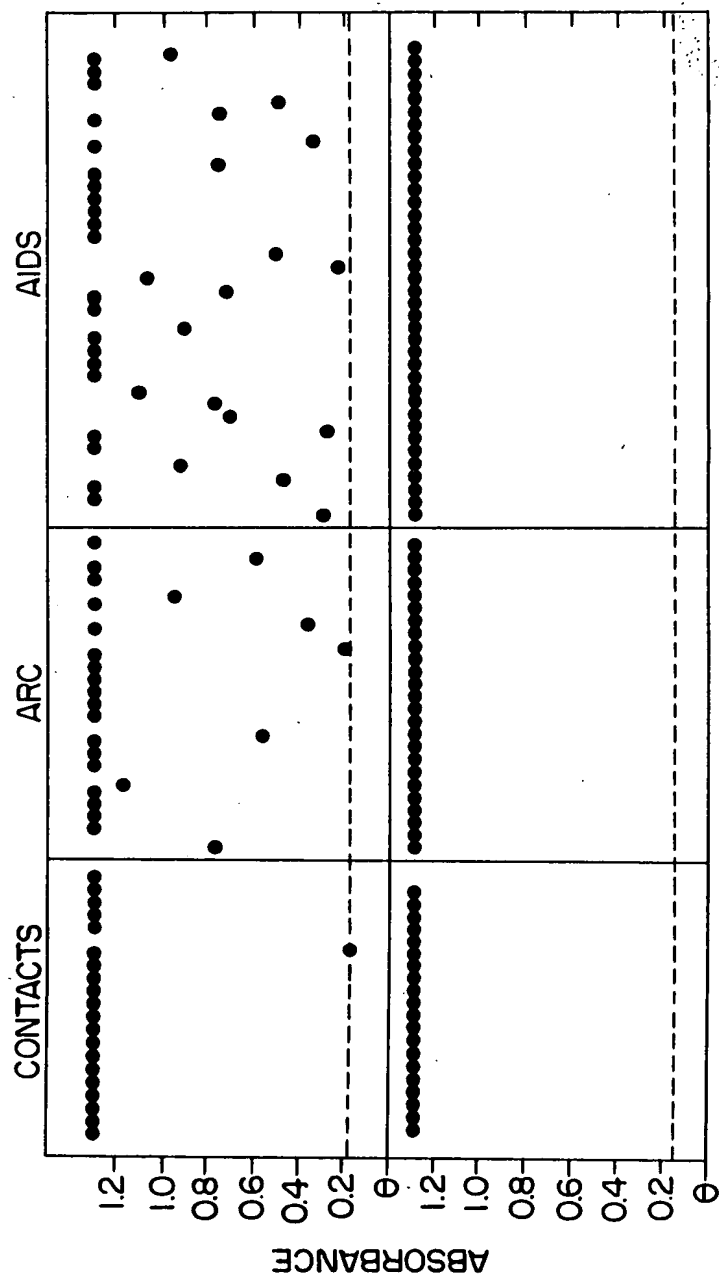


FIG.18

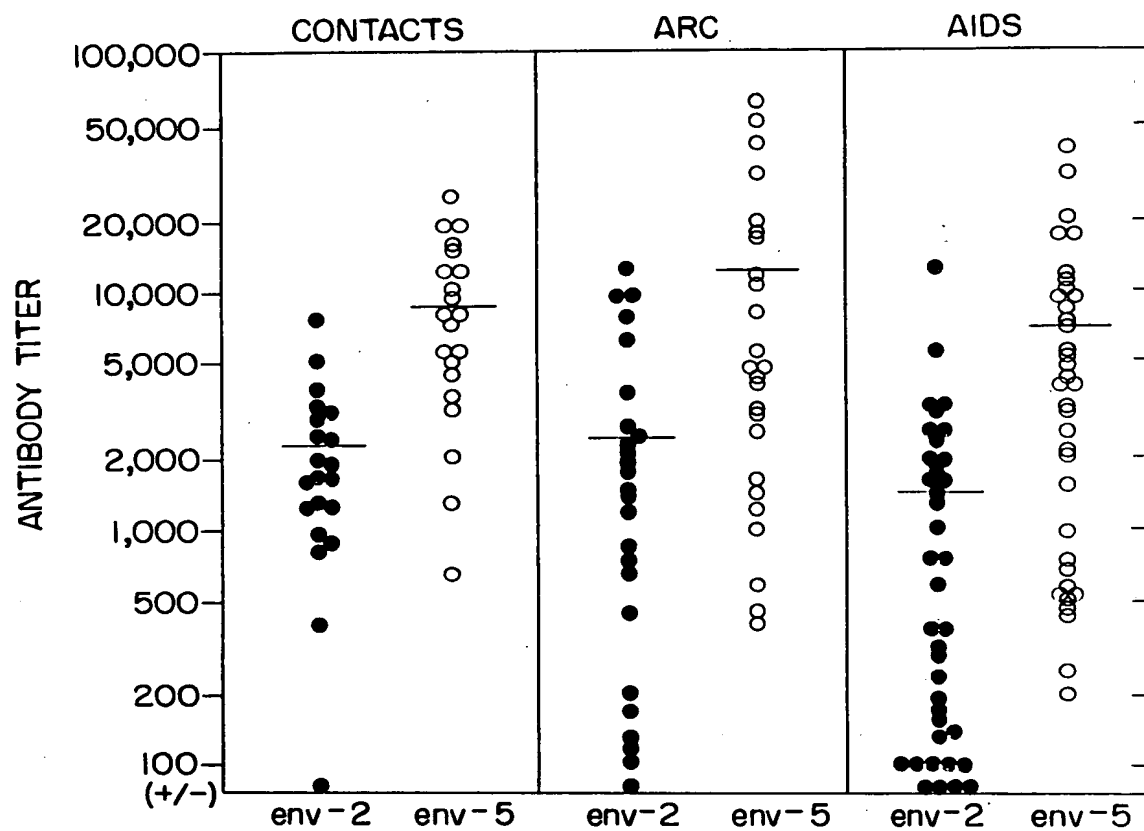


FIG.19

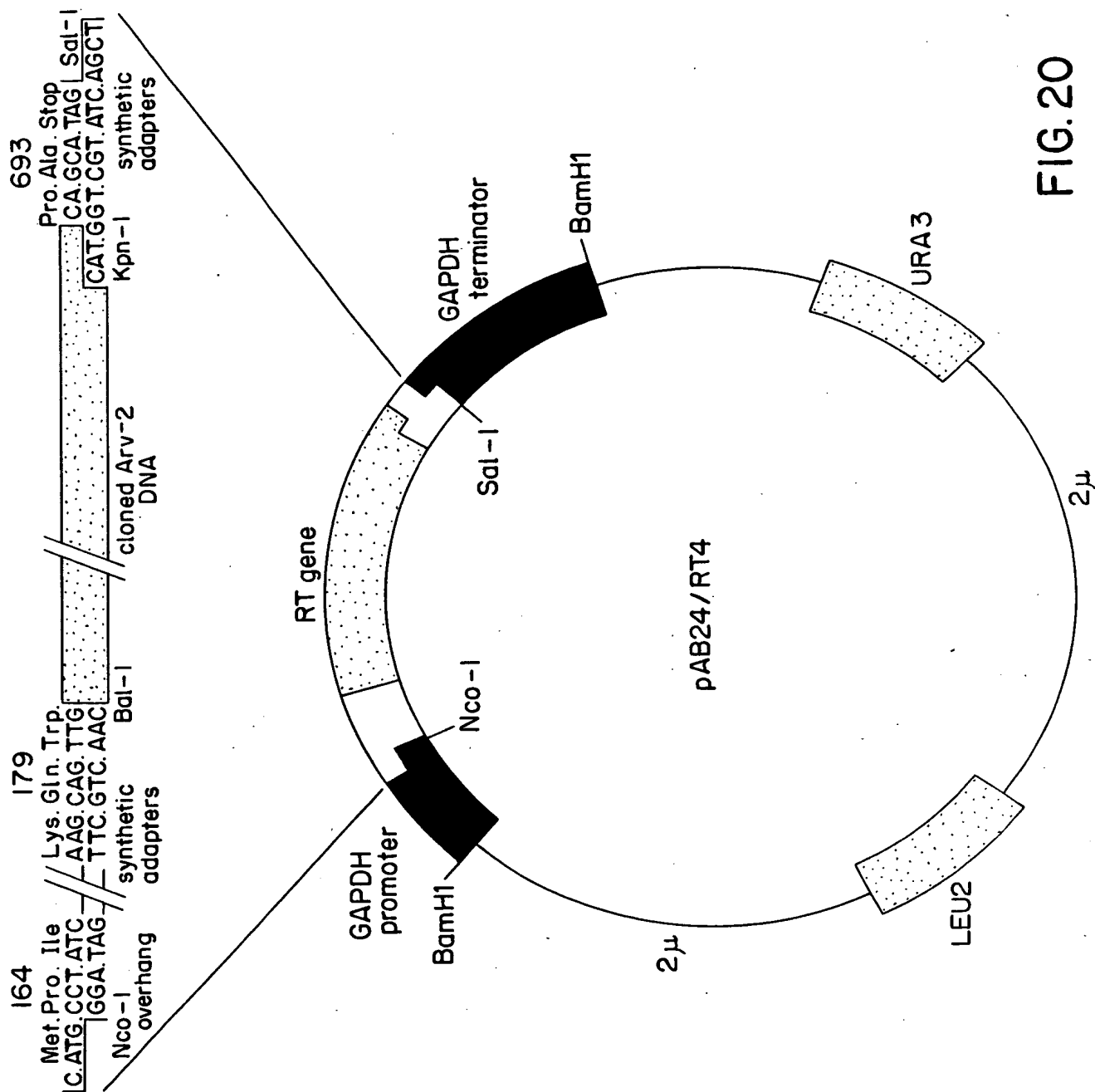


FIG.20

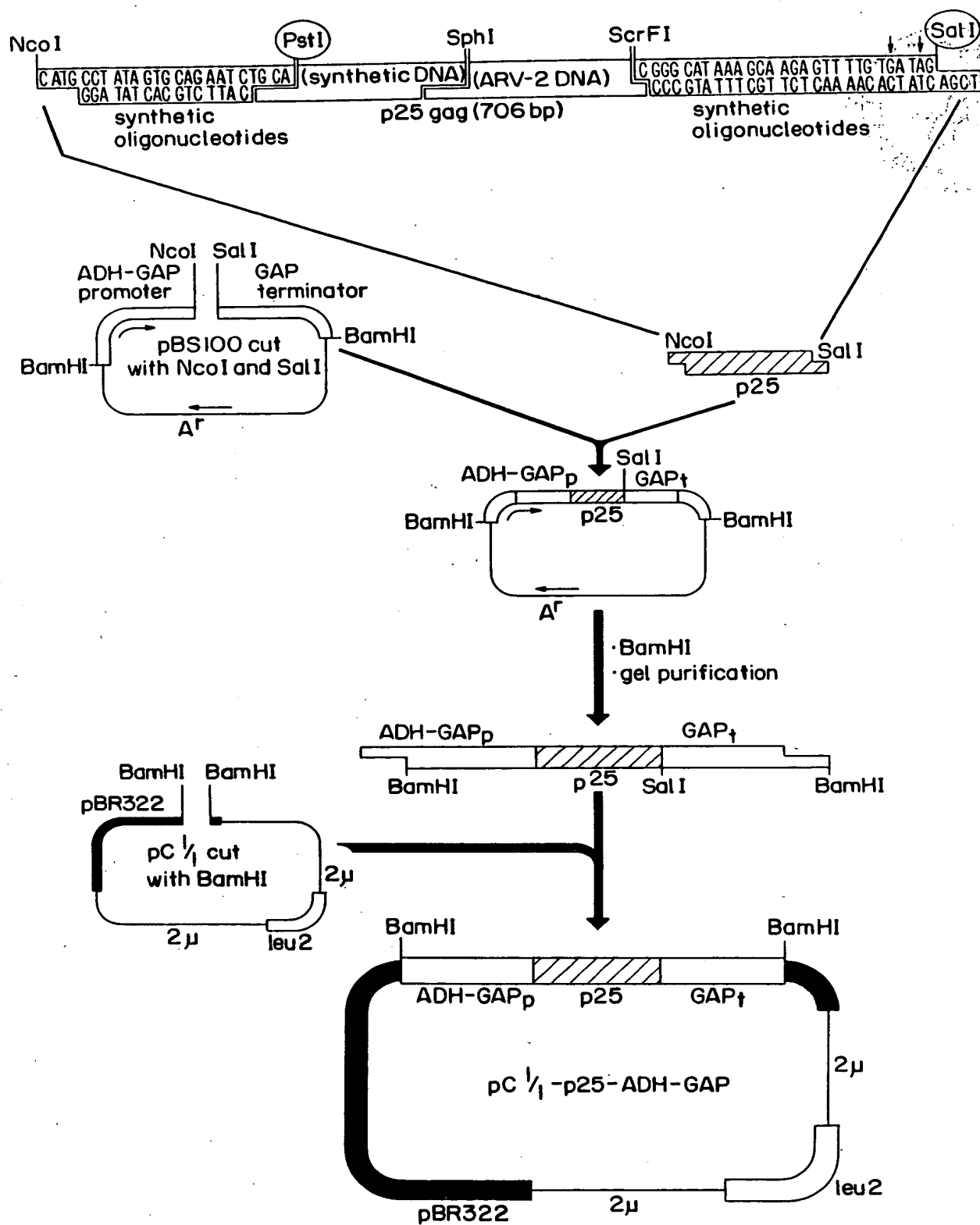


FIG.21

FIG. 22A

FIG. 22A

110
 Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met
 ATA GCA GGA ACT ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG
 120
 Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp
 ACA AAT AAT CCA CCT ATC CCA GTA GGA GAA ATC TAT AAA AGA TGG
 130
 Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr
 ATA ATC CTG GGA TTA AAT AAA ATA GTA AGA ATG TAT AGC CCT ACC
 140
 Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp
 AGC ATT CTG GAC ATA AGA CAA GGA CCA AAG GAA CCC TTT AGA GAT
 150
 Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln Ala Ser
 TAT GTA GAC CGG TTC TAT AAA ACT CTA AGA GCC GAA CAA GCT TCA
 160
 Gln Asp Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val Gln Asn
 CAG GAT GTA AAA AAT TGG ATG ACA GAA ACC TTG TTG GTC CAA AAT
 170
 Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Ala
 GCA AAC CCA GAT TGT AAG ACT ATT TTA AAA GCA TTG GGA CCA GCA
 180
 Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
 GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG GGA GTG GGG GGA
 190
 Pro Gly His Lys Ala Arg Val Leu OP
 CCC GGG CAT AAA GCA AGA GTT TTG TGA TAG

Translated Mol. Weight = 25700.75

FIG. 22B

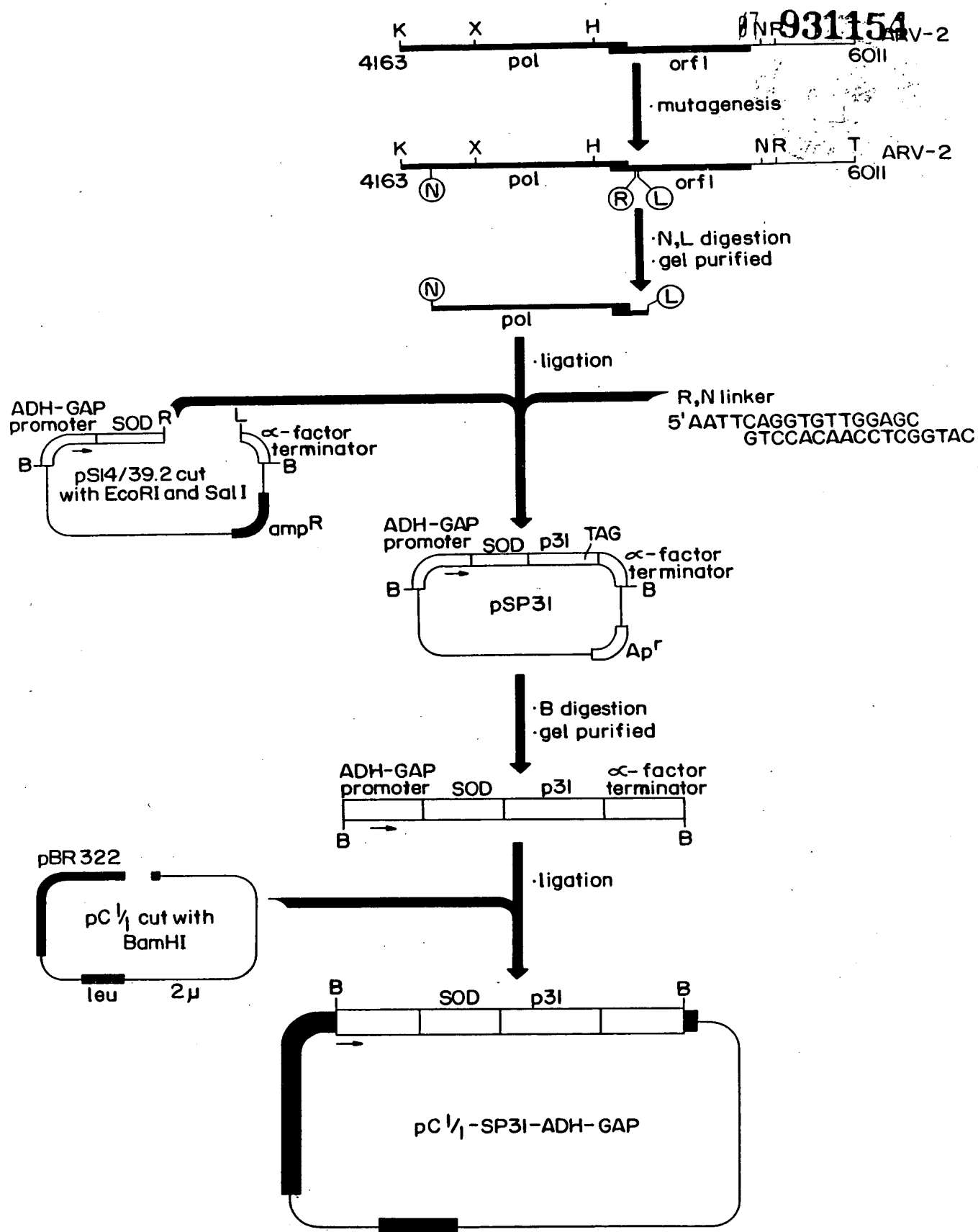


FIG. 23

SOD--▶
MetAlaThrLysAla
ATGGCTACAAAGGCT
TACCGATGTTCCGA

1383 ValCysValLeuLysGlyAspGlyProValGlnGlyIleIleAsnPheGluGlnLysGlu
GTTTGTGTTTTGAAGGGTGACGGCCAGTTCAAGGTATTATTAACCTTCGAGCAGAAGGAA
CAAACACAAAACCTTCCCACTGCCGGGTCAAGTTCATAATAATTGAAGCTCGTCTTCCTT

1443 SerAsnGlyProValLysValTrpGlySerIleLysGlyLeuThrGluGlyLeuHisGly
AGTAATGGACCAGTGAAAGGTGTGGGGAAGCATTAAAGGACTGACTGAAGGCCTGCATGGA
TCATTACCTGGTCACTTCCACACCCCTTCGTAATTTCTGACTGACTTCCGGACGTACCT

1503 PheHisValHisGluPheGlyAspAsnThrAlaGlyCysThrSerAlaGlyProHisPhe
TTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGTGCAGGTCCTCACTTT
AAGGTACAAGTACTCAAACCTCTATTATGTCGTCGACATGGTCACGTCCAGGAGTGAAA

1563 AsnProLeuSerArgLysHisGlyGlyProLysAspGluGluArgHisValGlyAspLeu
AATCCTCTATCCAGAAAACACGGTGGGCCAAAGGATGAAGAGAGGCATGTTGGAGACTTG
TTAGGAGATAGGTCTTTTGTGCCACCCGGTTTCTACTTCTCTCCGTACAACCTCTGAAC

1623 GlyAsnValThrAlaAspLysAspGlyValAlaAspValSerIleGluAspSerValIle
GGCAATGTGACTGCTGACAAAGATGGTGTGGCCGATGTGTCTATTGAAGATTCTGTGATC
CCGTTACACTGACGACTGTTTCTACCACACCCGGCTACACAGATAACTTCTAAGACACTAG

1683 SerLeuSerGlyAspHisCysIleIleGlyArgThrLeuValValHisGluLysAlaAsp
TCACTCTCAGGAGACCATTGCATCATTGGCCGCACACTGGTGGTCCATGAAAAAGCAGAT
AGTGAGAGTCCTCTGGTAACGTAGTAACCGGCGTGTGACCACCAGGTACTTTTTCGTCTA

1743 AspLeuGlyLysGlyGlyAsnGluGluSerThrLysThrGlyAsnAlaGlySerArgLeu
GACTTGGGCAAAGGTGGAAATGAAGAAAGTACAAAGACAGGAAACGCTGGAAGTCGTTTG
CTGAACCCGTTTCCACCTTTACTTCTTTTCATGTTTCTGTCTTTGCGACCTTCAGCAAAC

1803 AlaCysGlyValIleGlyIleAlaGlnAsnSerGlyValGlyAlaMetAlaMetAlaSer
GCTTGTGGTGTAATTGGGATCGCCAGAAATTCAGGTGTTGGAGCCATGGCCATGGCTAGT
CGAACACCACATTAACCCTAGCGGGTCTTAAGTCCACAACCTCGGTACCGGTACCGATCA

1863 AspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLysCysGln
GATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAG
CTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTGACACTATTTACAGTC

1923 LeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAsp
CTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGAT
GATTTTCCTCTTCGGTACGTACCTGTTTCATCTGACATCAGGTCCTTATACCGTTGATCTA

FIG. 24A

1983 CysThrHisLeuGluGlyLysIleIleLeuValAlaValHisValAlaSerGlyTyrIle
 TGTACACATCTAGAAGGAAAAATTATCCTGGTAGCAGTTTCATGTAGCCAGTGGATATATA
 ACATGTGTAGATCTTCCTTTTTAATAGGACCATCGTCAAGTACATCGGTCACCTATATAT

 2043 GluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeu
 GAAGCAGAAGTTATTCCAGCAGAGACAGGGCAGGAAACAGCATATTTTCTTTAAAATTAA
 CTTTCGTCTTCAATAAGGTCGTCTCTGTCCCGTCCTTTGTCGTATAAAAGAGAATTTTAAAT

 2103 AlaGlyArgTrpProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerThr
 GCAGGAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGCAATTTCCACAGTACT
 CGTCCTTCTACCGGTCATTTTTGTTATGTATGTCTGTTACCGTCGTTAAAGTGGTCATGA

 2163 ThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsn
 ACGGTTAAGGCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAAT
 TGCCAATTCCGGCGGACAACCAACCCGTCCTAGTTGTCCTTAAACCGTAAGGGATGTTA

 2223 ProGlnSerGlnGlyValValGluSerMetAsnAsnGluLeuLysLysIleIleGlyGln
 CCCCAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAGAAAAATTATAGGACAG
 GGGGTTTCAGTTCTCATCATCTTAGATACTTATTACTTAATTTCTTTTAATATCCTGTC

 2283 ValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIleHisAsn
 GTAAGAGATCAGGCTGAACACCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAAT
 CATTCTCTAGTCCGACTTGTGGAATTCTGTCGTCATGTTTACCGTCATAAGTAGGTGTTA

 2343 PheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleIle
 TTTAAAAGAAAAGGGGGGATTGGGGGATACAGTGCAGGGGAAAGAAATAGTAGACATAATA
 AAATTTTCTTTTCCCCCTAACCCCTATGTCACGTCCCTTTCTTATCATCTGTATTAT

 2403 AlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArg
 GCAACAGACATACAACTAAAGAACTACAAAAGCAAATTACAAAAATTCAAAATTTTCGG
 CGTTGTCGTATGTTTGATTTCTTGATGTTTTCGTTTAATGTTTTTAAGTTTTAAAGCC

 2463 ValTyrTyrArgAspAsnLysAspProLeuTrpLysGlyProAlaLysLeuLeuTrpLys
 GTTTATTACAGGGACAACAAAGATCCCTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAA
 CAAATAATGTCCCTGTTGTTTCTAGGGGAAACCTTTCCTGGTCGTTTCGAAGAGACCTTT

 2523 GlyGluGlyAlaValValIleGlnAspAsnSerAspIleLysValValProArgArgLys
 GGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAA
 CCACTTCCCCGTCATCATTATGTTCTATTATCACTGTATTTTCATCACGGTTCTTCTTTT

 2583 AlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArg
 GCAAAAATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGA
 CGTTTTTAGTAATCCCTAATACCTTTTGTCTACCGTCCACTACTAACACACCGTTTCATCT

 2643 GlnAspGluAspAM
 CAGGATGAGGATTAG
 GTCCTACTCCTAATC

FIG. 24B

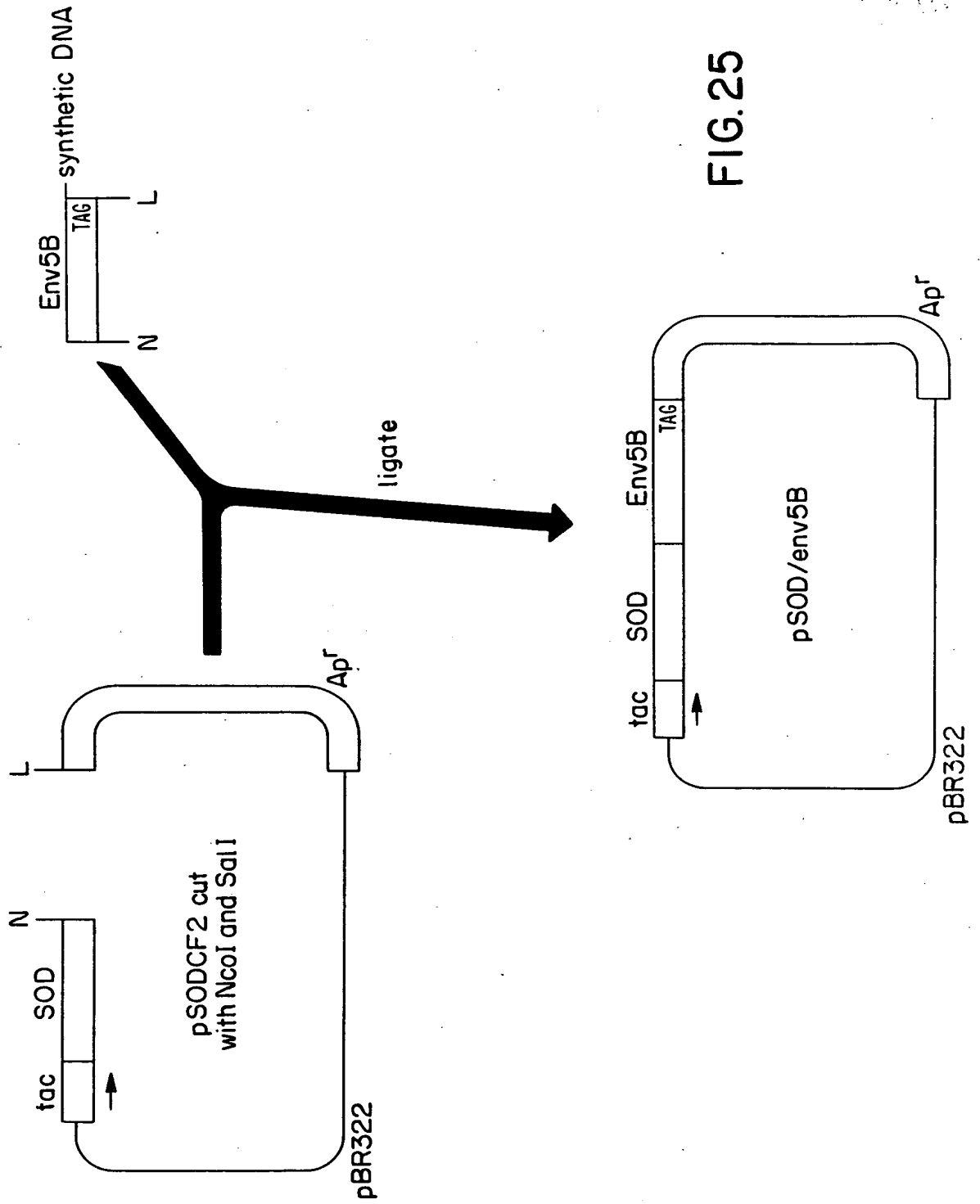


FIG.25

Sequence of SOD/env-4

SOD --►
 MetAlaThrLysAlaValCysValLeuLysGlyAspGlyProValGlnGlyIleIleAsn
 1 CATGGCGACGAAGGCCGTGTGCGTGCTGAAGGGCGACGGCCAGTGCAGGGCATCATCAAT
 CGCTGCTTCCGGCACACGCACGACTTCCCCTGCCGGGTACAGTCCCGTAGTAGTTA

62 PheGluGlnLysGluSerAsnGlyProValLysValTrpGlySerIleLysGlyLeuThr
 TTCGAGCAGAAGGAAAGTAATGGACCAGTGAAGGTGTGGGGAAGCATTAAAGGACTGACT
 AAGCTCGTCTTCTTTTATTACCTGGTCACTTCCACACCCCTTCGTAATTTCTGACTGA

122 GluGlyLeuHisGlyPheHisValHisGluPheGlyAspAsnThrAlaGlyCysThrSer
 GAAGGCCTGCATGGATTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGT
 CTTCCGGACGTACCTAAGGTACAAGTACTCAAACCTCTATTATGTCGTCCGACATGGTCA

182 AlaGlyProHisPheAsnProLeuSerArgLysHisGlyGlyProLysAspGluGluArg
 GCAGGTCTCTCACTTTAATCCTCTATCCAGAAAACACGGTGGGCCAAAGGATGAAGAGAGG
 CGTCCAGGAGTGAAATTAGGAGATAGGTCTTTTGTGCCACCCGGTTTCTACTTCTCTCC

242 HisValGlyAspLeuGlyAsnValThrAlaAspLysAspGlyValAlaAspValSerIle
 CATGTTGGAGACTTGGGCAATGTGACTGCTGACAAAGATGGTGTGGCCGATGTGTCTATT
 GTACAACCTCTGAACCCGTTACACTGACGACTGTTTCTACCACACCCGGCTACACAGATAA

302 GluAspSerValIleSerLeuSerGlyAspHisCysIleIleGlyArgThrLeuValVal
 GAAGATTCTGTGATCTCACTCTCAGGAGACCATTGCATCATTGGCCGCACACTGGTGGTC
 CTTCTAAGACACTAGAGTGAGAGTCTCTGGTAACGTAGTAACCGGCGTGTGACCACCAG

362 HisGluLysAlaAspAspLeuGlyLysGlyGlyAsnGluGluSerThrLysThrGlyAsn
 CATGAAAAAGCAGATGACTTGGGCAAAGGTGGAAATGAAGAAAGTACAAAGACAGGAAAC
 GTACTTTTTCTGCTACTGAACCCGTTTCCACCTTTACTTCTTTTCATGTTTCTGTCCTTTG

Env4--►
 422 AlaGlySerArgLeuAlaCysGlyValIleGlnIleAlaMetGluValValIleArgSer
 GCTGGAAGTCGTTTGGCTTGTGGTGTAAATTGGGATCGCCATGGAGGTAGTAATTAGATCT
 CGACCTTCAGCAAACCGAACACCACATTAACCTAGCGGTACCTCCATCATTAACTAGA

482 AspAsnPheThrAsnAsnAlaLysThrIleIleValGlnLeuAsnGluSerValAlaIle
 GACAATTTACGAACAATGCTAAAACCATAATAGTACAGCTGAATGAATCTGTAGCAATT
 CTGTTAAAGTGCTTGTACGATTTTGGTATTATCATGTCGACTTACTTAGACATCGTTAA

542 AsnCysThrArgProAsnAsnAsnThrArgLysSerIleTyrIleGlyProGlyArgAla
 AACTGTACAAGACCCAACAACAATACAAGAAAAAGTATCTATATAGGACCAGGGAGAGCA
 TTGACATGTTCTGGGTGTTGTTATGTTCTTTTTCATAGATATATCCTGGTCCCTCTCGT

FIG. 26A

602 PheHisThrThrGlyArgIleIleGlyAspIleArgLysAlaHisCysAsnIleSerArg
TTTCATACAACAGGAAGAATAATAGGAGATATAAGAAAAGCACATTGTAACATTAGTAGA
AAAGTATGTTGTCCTTCTTATTATCCTCTATATTCTTTTCGTGTAACATTGTAATCATCT

662 AlaGlnTrpAsnAsnThrLeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsn
GCACAATGGAATAACACTTTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGGAAT
CGTGTTACCTTATTGTGAAATCTTGCTATCAATTTTTTAATTCTCTTGTCAAACCCTTA

722 AsnLysThrIleValPheAsnGlnSerSerGlyGlyAspProGluIleValMetHisSer
AATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGT
TTATTTTGTATCAGAAATTAGTTAGGAGTCCTCCCCTGGGTCTTTAACATTACGTGTCA

782 PheAsnCysArgGlyGluPhePheTyrCysAsnThrThrGlnLeuPheAsnAsnThrTrp
TTTAATTGTAGAGGGGAATTTTTCTACTGTAATACAACACAACCTGTTTAATAATACATGG
AAATTAACATCTCCCCTTAAAAAGATGACATTATGTTGTGTTGACAAATTATTATGTACC

842 ArgLeuAsnHisThrGluGlyThrLysGlyAsnAspThrIleIleLeuProCysArgIle
AGGTTAAATCACACTGAAGGAACTAAAGGAAATGACACAATCATACTCCCATGTAGAATA
TCCAATTTAGTGTGACTTCCTTGATTTCTTTACTGTGTTAGTATGAGGGTACATCTTAT

902 LysGlnIleIleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGly
AAACAAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATTGGA
TTTGTTTAATATTTGTACACCGTCCTTCATCTTTTCGTTACATACGGGGAGGGTAACCT

962 GlyGlnIleSerCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThr
GGACAAATTAGTTGTTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTACA
CCTGTTTAATCAACAAGTAGTTTATAATGTCCCGACGATAATTGTTCTCTACCACCATGT

1022 AsnValThrAsnAspThrGluValPheArgProGlyGlyGlyAspMetArgAspAsnTrp
AATGTAACATAATGACACCGAGGTCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGG
TTACATTGATTACTGTGGCTCCAGAAGTCTGGACCTCCTCCTCTATACTCCCTGTAAACC

1082 ArgSerGluLeuTyrLysTyrLysValIleLysIleGluProLeuGlyIleAlaProThr
AGAAGTGAATTATATAAATATAAAGTAATAAAAAATTGAACCATTAGGAATAGCACCACC
TCTTCACTTAATATATTTATATTTTCAATTATTTTAACTTGGTAATCCTTATCGTGGGTGG

1142 LysAlaLysArgArgValValGlnArgGluLysArgOP OP
AAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGATGATGAAGCTTG
TTCCGTTTCTTCTCACCACGTCTCTTTTTTCTACTACTTCGAACAGCT

FIG. 26B

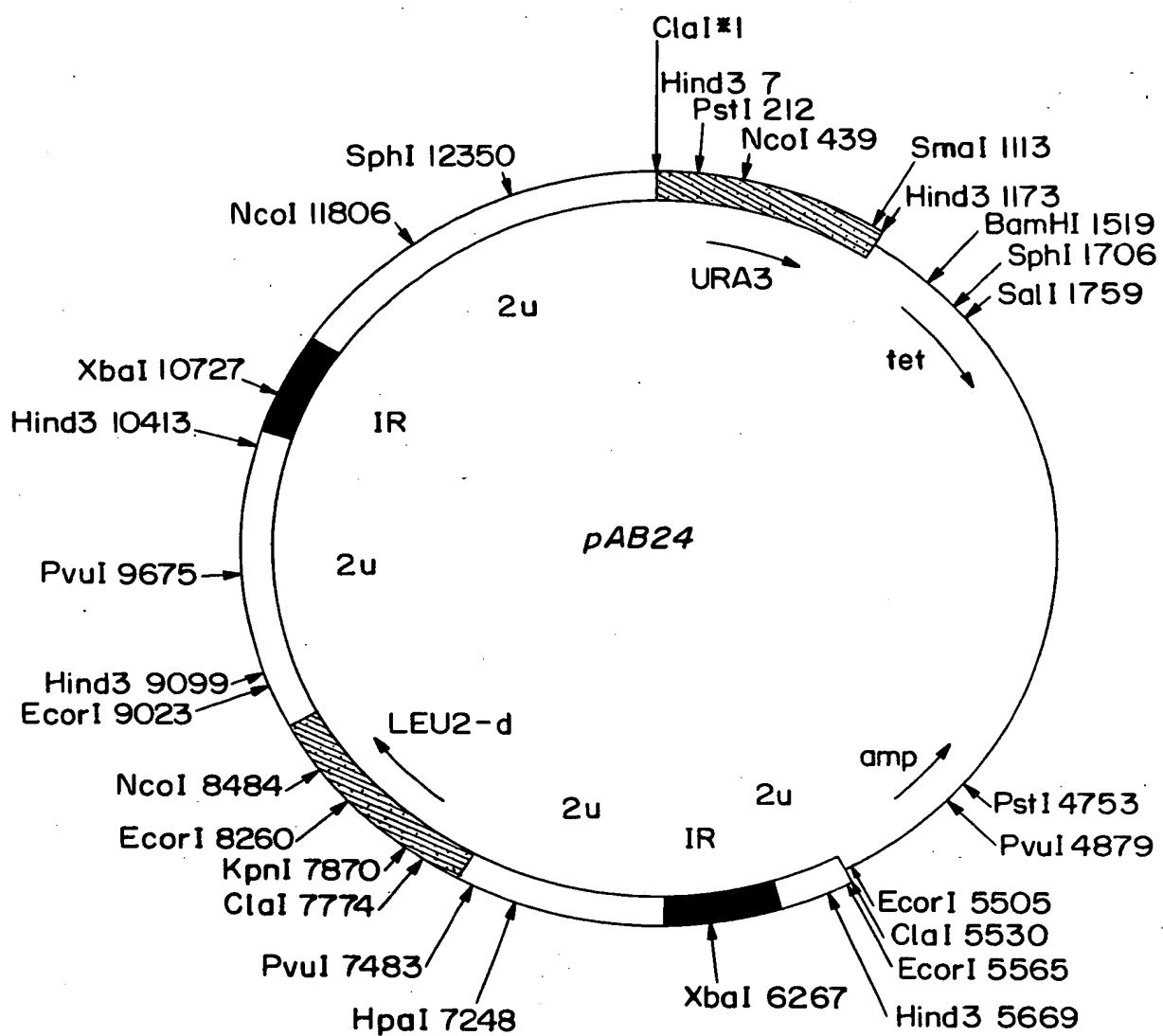


FIG. 27

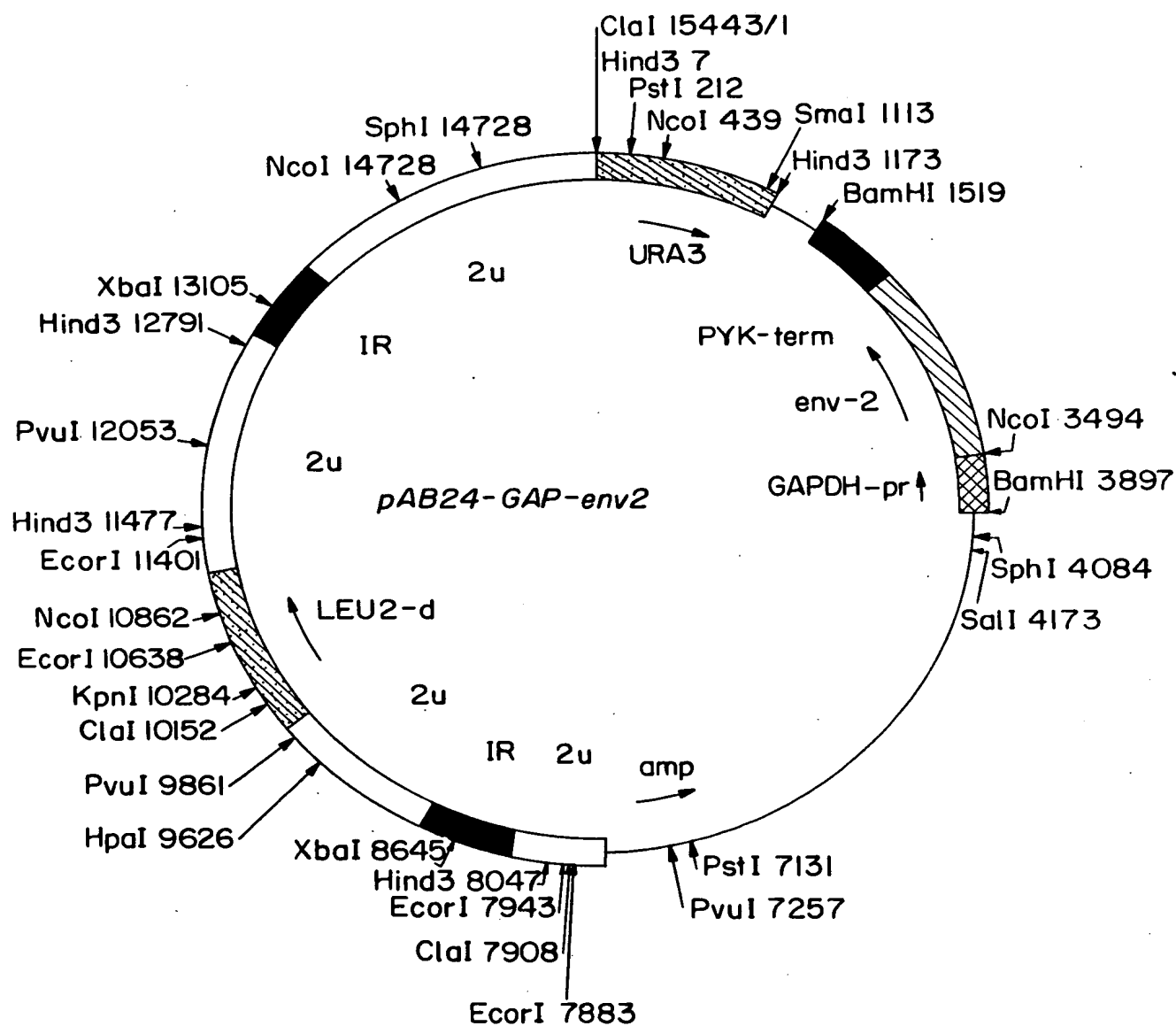


FIG. 28

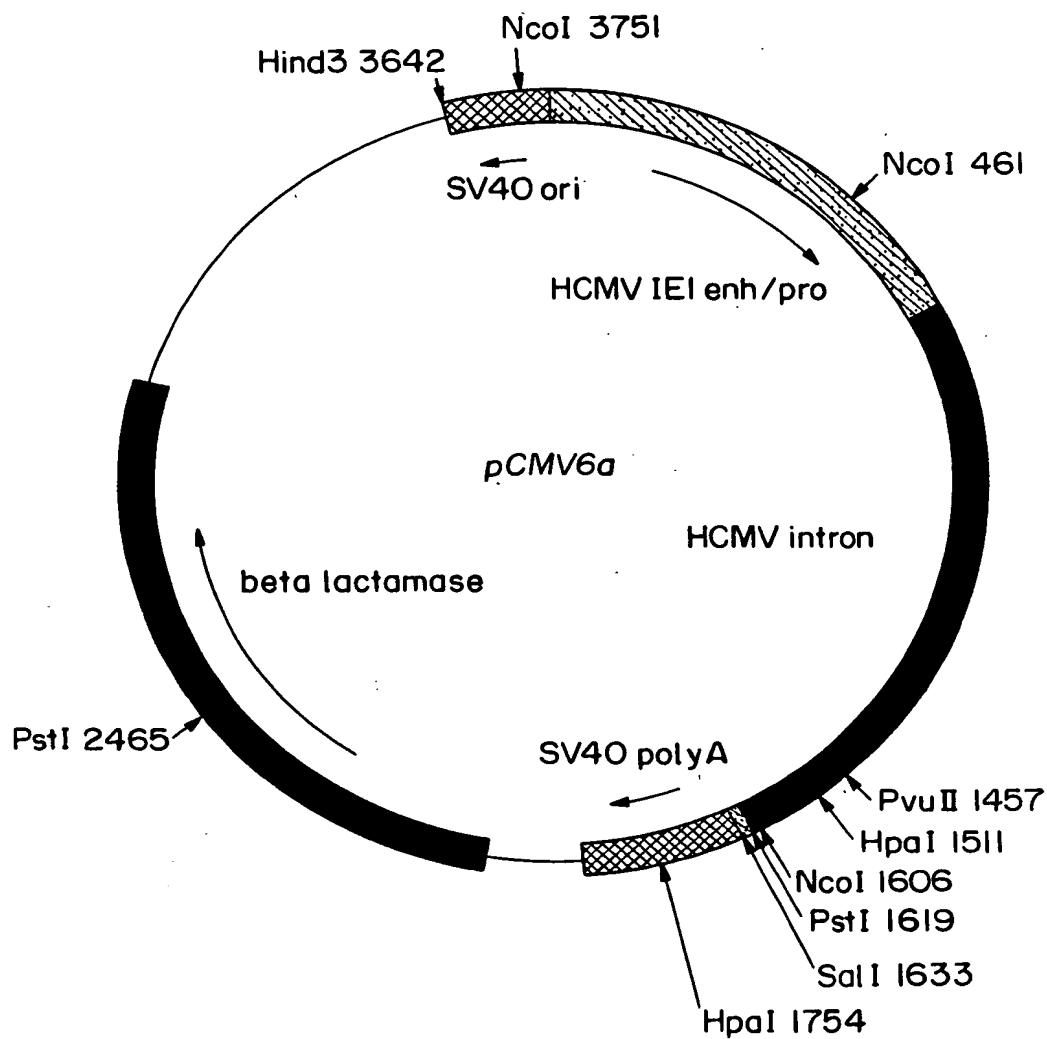


FIG. 29

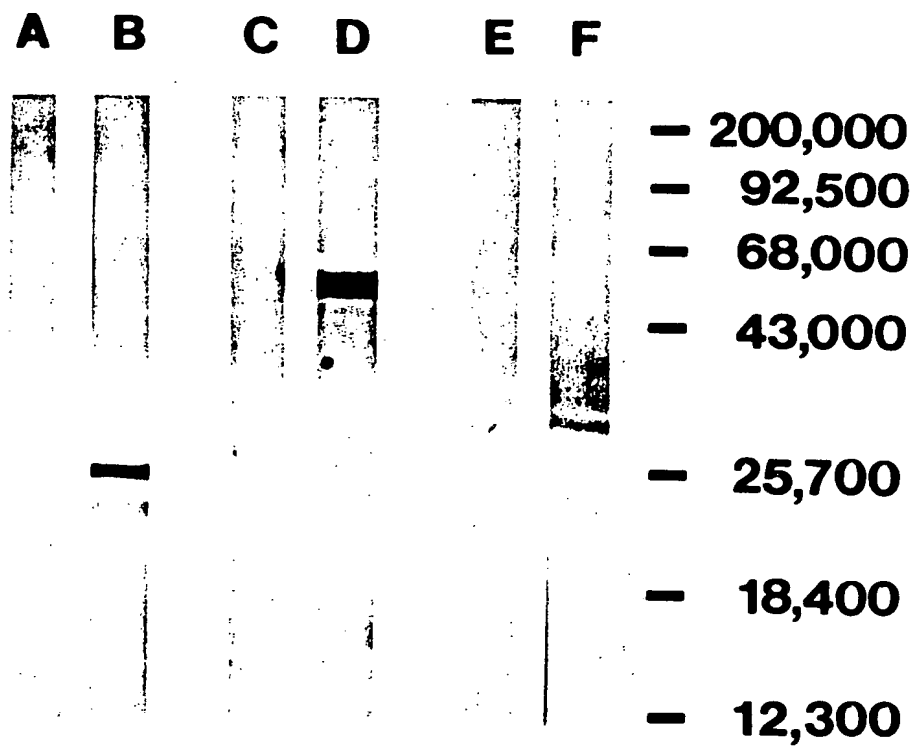


FIG. 30

FIG. 3la

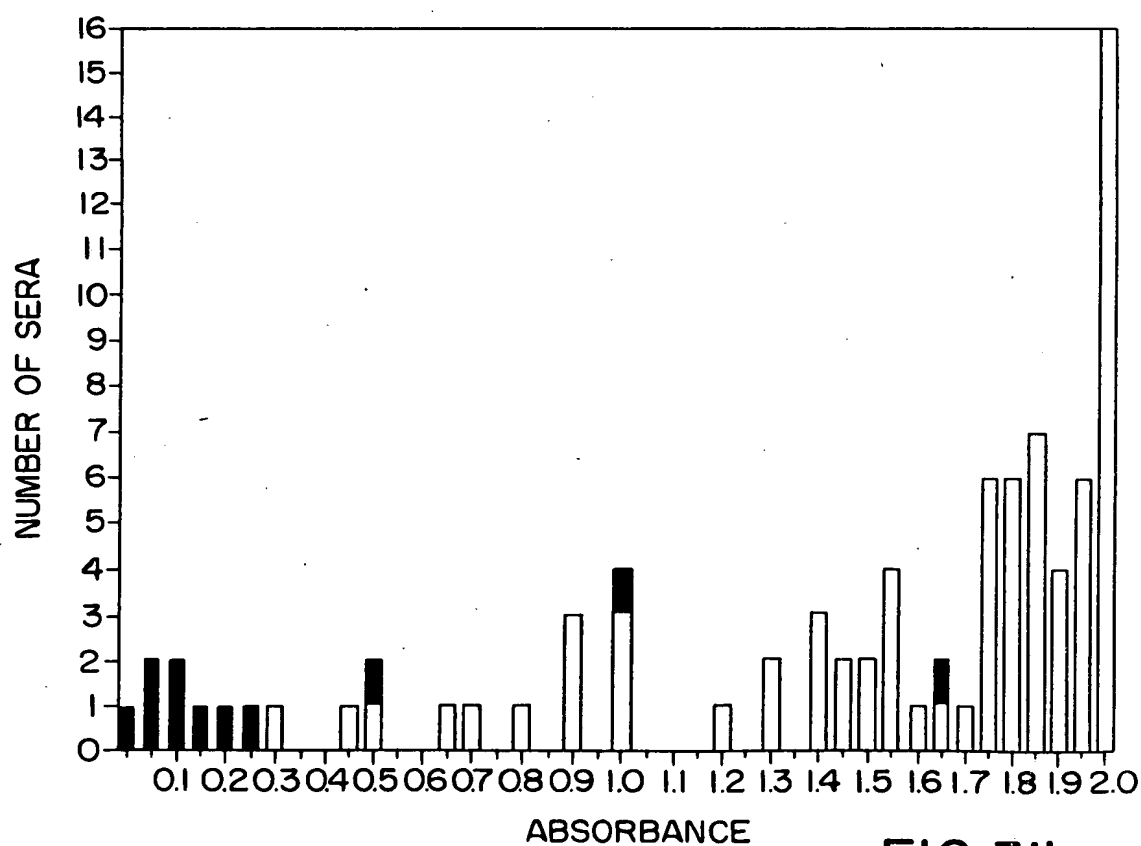
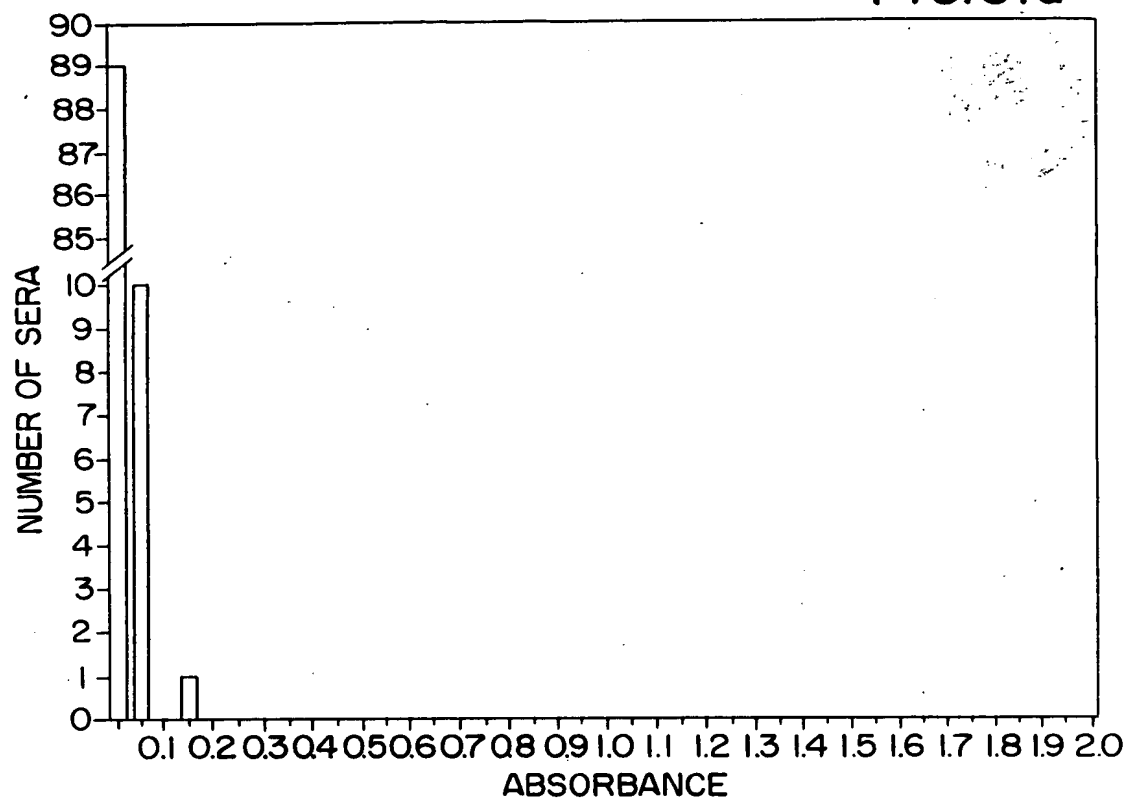


FIG. 3lb